Indiana Law Review

Volume 37 No. 1 2003

Indiana Law Review

Volume 37 2003-2004

INDIANA LAW REVIEW

(ISSN 0090-4198)

Published four times a year by Indiana University. Editorial and business Offices are located at:

Indiana Law Review
Lawrence W. Inlow Hall
530 W. New York Street
Indianapolis, IN 46202-3225
(317) 274-4440

Subscriptions. Current subscription rates for an academic year are $30.00 (domestic mailing) and $35.00 (foreign mailing) for four issues. Unless the Business Office receives notice to the contrary, all subscriptions will be renewed automatically. ***Address changes must be received at least one month prior to publication to ensure prompt delivery and must include old and new address and the proper zip code.***

Single and Back Issues. Current single regular issues may be obtained from the *Indiana Law Review* Business Office for $10.00. Current Survey Issues may be purchased for $20.00. The Symposium Issue may be purchased for $15.00. Please enclose payment with order. Claims for nonreceipt of current year's issues must be made within six months of the mailing date. Back issues may be ordered from: William S. Hein & Co., Inc., 1285 Main St., Buffalo, NY 14209-1987, (800) 828-7571, (716) 883-8100 fax, mail@wshein.com (e-mail for orders).

Manuscripts. The *Indiana Law Review* invites submission of unsolicited manuscripts. All citations should conform to *The Bluebook A Uniform System of Citation* (17th ed. 2000). Manuscripts should be double-spaces with standard one-inch margins. The *Indiana Law Review*, as a matter of policy, promotes the use of gender-neutral language, and alternate use of the female or male pronoun is intended to refer to both female and male.

The INDIANA LAW REVIEW (ISSN 0090-4198) is the property of Indiana University and is published quarterly by the Indiana University School of Law—Indianapolis, which assumes complete editorial responsibility thereof. Current subscription rates for an academic year are: one year domestic, $30.00; foreign, $35.00. Send all correspondence to Editorial Specialist, *Indiana Law Review*, Indiana University School of Law—Indianapolis, Lawrence W. Inlow Hall, 530 W. New York Street, Indianapolis, Indiana 46202-3225. Publication office: 530 W. New York Street, Indianapolis, Indiana 46202-3225. Periodicals paid at Indianapolis, Indiana 46201.

POSTMASTER: Send address changes to INDIANA LAW REVIEW, Lawrence W. Inlow Hall, 530 W. New York Street, Indianapolis, Indiana 46202-3225.

The entire text of this Law Review is printed on recycled paper.

Indiana International & Comparative Law Review

The *Indiana International & Comparative Law Review*, published by the Indiana University School of Law- Indianapolis since 1991, is a student-edited law journal that provides a forum for the discussion and analysis of contemporary issues in public and private international law. The *Review* publishes articles by prominent legal scholars, practitioners, and policy makers from around the world, as well as student-written notes and comments.

The *Review* is published twice annually, with a periodic third issue devoted to a specific topic in international or comparative law. In 1994, in conjunction with the Italian Academy for Advanced Studies in America at Columbia University, the *Review* published a special issue on the political and social aspects of Italian law. In 1996, the *Review* published an issue addressing various aspects of Chinese law.

Name

School/Firm/Business

Address

City, State, Zip Code

Subscriptions within the U.S. are $18.00 per year and outside the U.S. are $21.00 per year. I enclose $ _____ for _____ subscription(s) to the *Indiana International & Comparative Law Review.* Please make your check payable to *Indiana International & Comparative Law Review.*

Mail to: Executive Production Editor
Indiana International & Comparative Law Review
Lawrence W. Inlow Hall
530 West New York Street
Indianapolis, Indiana 46202

Please enter my subscription to the INDIANA LAW REVIEW

NAME _____

ADDRESS: _____

_____

_____

Enclosed is $ _____ for _____ subscription(s)

Mail to: **ATTN: Law Review Specialist**
at
INDIANA LAW REVIEW
INDIANA UNIVERSITY SCHOOL OF LAW—**INDIANAPOLIS**
Lawrence W. Inlow Hall
530 West New York Street
Indianapolis, Indiana 46202-3225

For an *academic* year, the subscription rate for four issues is:
Domestic, $30; Foreign, $35; Student, $20
Single Issue, $10; Survey Issue, $20
Symposium Issue, $15

Indiana University School of Law—Indianapolis

2002-2003 ADMINISTRATIVE OFFICERS AND FACULTY

Administrative Officers

ADAM W. HERBERT, *President of the University.* B.A., M.A., University of Southern California; Ph.D., University of Pittsburgh.

CHARLES R. BANTZ, *Chancellor, Indiana University-Purdue University—Indianapolis.* B.A., M.A., University of Minnesota; Ph.D., The Ohio State University.

ANTHONY A. TARR, *Dean and Professor of Law.* B.A., LL.B., University of Natal; LL.M., Cambridge University; Ph.D., University of Canterbury; Ph.D., Cambridge University.

THOMAS B. ALLINGTON, *Associate Dean for Technology and Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University.

JEFFREY W. GROVE, *Associate Dean for Graduate Studies, Professor of Law, and Director, China Summer Program.* A.B., Juniata College; J.D., George Washington University Law School.

SUSANAH M. MEAD, *Associate Dean for Academic Affairs and Professor of Law.* B.A., Smith College; J.D., Indiana University School of Law—Indianapolis.

CYNTHIA BAKER, *Director, Program on Law and State Government.* B.A., J.D., Valparaiso University.

ELIZABETH L. DECOUX, *Assistant Dean for Student Affairs.* J.D., Mississippi College School of Law.

ANGELA M. ESPADA, *Assistant Dean for Admissions.* J.D., Indiana University School of Law—Bloomington.

JONNA KANE MACDOUGALL, *Assistant Dean for External Affairs.* J.D., Indiana University School of Law—Indianapolis.

CAROL B. NEARY, *Director of Development.* B.A., Indiana University.

SHANNON L. WILLIAMS, *Director of Professional Development.* B.S., Indiana University.

Faculty

CYNTHIA M. ADAMS, *Clinical Associate Professor of Law.* B.A., Kentucky Wesleyan College; J.D., Indiana University School of Law—Indianapolis.

THOMAS B. ALLINGTON, *Associate Dean for Technology and Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University School of Law.

JUDITH FORD ANSPACH, *Professor of Law and Director, Ruth Lilly Law Library.* B.S., M.L.S., Kent State University; J.D., Mississippi College School of Law.

GERALD L. BEPKO, *Indiana University-Purdue University—Indianapolis Chancellor Emeritus, Indiana University Trustee Professor and Professor of Law.* B.S., Northern Illinois University; J.D., ITT/Chicago-Kent College of Law; LL.M., Yale Law School.

FRANK BOWMAN, *M. Dale Palmer Professor of Law.* B.A., Colorado College; J.D., Harvard Law School.

WILLIAM C. BRADFORD, *Assistant Professor of Law.* B.A., M.A., University of Miami; Ph.D., Northwestern University; J.D., University of Miami School of Law; LL.M., Harvard Law School.

ROBERT BROOKINS, *Professor of Law.* B.S., University of South Florida; J.D., Ph.D., Cornell University.

KENNETH D. CHESTEK, *Clinical Associate Professor of Law.* B.A., Pennsylvania State University; J.D., University of Pittsburgh School of Law.

DANIEL H. COLE, *R. Bruce Townsend Professor of Law.* A.B., Occidental College; A.M., University of Chicago; J.D., Lewis and Clark Law School; J.S.M., J.S.D., Stanford Law School.

JEFFREY O. COOPER, *Associate Professor of Law.* A.B., Harvard University; J.D., University of Pennsylvania Law School.

PAUL N. COX, *Centennial Professor of Law.* B.S., Utah State University; J.D., University of Utah College of Law; LL.M., University of Virginia School of Law.

ROBIN KUNDIS CRAIG, *Associate Professor of Law.* B.A., Pomona College; M.A., The Johns Hopkins University; Ph.D., University of California; J.D., Lewis and Clark Law School.

KENNETH D. CREWS, *Associate Dean of the Faculties for Copyright Management and Samuel R. Rosen II Professor of Law and Professor of Library and Information Science.* B.A., Northwestern University; J.D., Washington University School of Law; M.L.S., University of California, Los Angeles; Ph.D., University of California, Los Angeles.

JAMES D. DIMITRI, *Clinical Associate Professor of Law.* B.S., Indiana University; J.D., Valparaiso University School of Law.

JENNIFER ANN DROBAC, *Associate Professor of Law.* B.A., M.A., Stanford University; J.D., J.S.D., Stanford Law School.

GEORGE E. EDWARDS, *Professor of Law and Director, Program in International Human Rights Law.* B.A., North Carolina State University; J.D., Harvard Law School.

FRANK EMMERT, *Professor of Law and Executive Director, Center for International and Comparative Law.* Erstes Juristisches Staatsexamen (J.D.), University of Munich Law School; LL.M., The University of Michigan Law School; Ph.D., University of Maastricht; Diploma, European University Institute.

NICHOLAS GEORGAKOPOULOS, *Professor of Law.* Ptyhion Nomikis, Athens University School of Law; LL.M., S.J.D., Harvard Law School.

JEFFREY W. GROVE, *Associate Dean, Professor of Law, and Director, China Summer Program.* A.B., Juniata College; J.D., George Washington University Law School.

FRANCES WATSON HARDY, *Clinical Associate Professor of Law.* B.S., Ball State University; J.D., Indiana University School of Law—Indianapolis.

JOHN LAWRENCE HILL, *Professor of Law.* B.A., Northern Illinois University; J.D., Ph.D., Georgetown University.

LAWRENCE A. JEGEN, III, *Thomas F. Sheehan Professor of Tax Law and Policy.* A.B., Beloit College; J.D., M.B.A., The University of Michigan; LL.M., New York University School of Law.

HENRY C. KARLSON, *Professor of Law.* A.B., J.D., LL.M., University of Illinois College of Law.

ROBERT A. KATZ, *Associate Professor of Law.* A.B., Harvard College; J.D., University of Chicago Law School.

LINDA KELLY-HILL, *Professor of Law.* B.A., J.D., University of Virginia.

ELEANOR D. KINNEY, *Hall Render Professor of Law and Co-Director of the Center for Law and Health.* B.A., Duke University; M.A., University of Chicago; J.D., Duke University School of Law; M.P.H., University of North Carolina.

ANDREW R. KLEIN, *Paul E. Beam Professor of Law.* B.A., University of Wisconsin; J.D., Emory University School of Law.

ROBERT E. LANCASTER, *Clinical Associate Professor of Law.* B.A., Millsaps College; J.D., Tulane Law School.

NORMAN LEFSTEIN, *Professor of Law and Dean Emeritus.* LL.B., University of Illinois College of Law; LL.M., Georgetown University Law School.

MARÍA PABÓN LÓPEZ, *Assistant Professor of Law and Co-Director, Latin American Law Program.* B.A., Princeton University; J.D., University of Pennsylvania Law School.

GERARD N. MAGLIOCCA, *Assistant Professor of Law.* B.A., Stanford University; J.D., Yale Law School.

ALLISON MARTIN, *Clinical Associate Professor Law.* B.S., J.D., University of Illinois.

DEBORAH MCGREGOR, *Clinical Associate Professor of Law and Assistant Director of Legal Analysis, Research and Communication.* B.A., University of Evansville; J.D., Georgetown University Law School.

SUSANAH M. MEAD, *Associate Dean and Professor of Law.* B.A., Smith College; J.D., Indiana University School of Law—Indianapolis.

MARY H. MITCHELL, *Professor of Law.* A.B., Butler University; J.D., Cornell Law School.

NOVELLA NEDEFF, *Clinical Associate Professor of Law.* B.A., J.D., Indiana University.

JAMES P. NEHF, *Cleon H. Foust Fellow, Professor of Law, and Director, European Law Program.* B.A., Knox College; J.D., University of North Carolina Law School.

DAVID ORENTLICHER, *Samuel R. Rosen I Professor of Law and Co-Director of the Center for Law and Health.* A.B., Brandeis University; J.D., M.D., Harvard College.

JOANNE ORR, *Clinical Associate Professor of Law.* B.S., Indiana State University; J.D., California Western School of Law.

ANTONY PAGE, *Assistant Professor of Law.* B. Comm., McGill University; M.B.A., Simon Fraser University; J.D., Stanford Law School.

H. KATHLEEN PATCHEL, *Associate Professor of Law.* A.B., Huntington College; J.D., University of North Carolina Law School; LL.M., Yale Law School.

FLORENCE WAGMAN ROISMAN, *Michael D. McCormick I Professor of Law.* B.A., University of Connecticut; LL.B., Harvard Law School.

JOAN M. RUHTENBERG, *Clinical Professor of Law and Director of Legal Analysis, Research and Communication.* B.A., Mississippi University for Women; J.D., Indiana University School of Law—Indianapolis.

JOEL M. SCHUMM, *Clinical Assistant Professor of Law.* B.A., Ohio Wesleyan University; M.A., University of Cincinnati; J.D., Indiana University School of Law—Indianapolis.

ANTHONY A. TARR, *Dean and Professor of Law.* B.A., LL.B., University of Natal; LL.M., Cambridge University; Ph.D., University of Canterbury; Ph.D., Cambridge University.

JAMES W. TORKE, *Carl M. Gray Professor of Law.* B.S., J.D., University of Wisconsin.

LAWRENCE P. WILKINS, *William R. Neale Professor of Law.* B.A., The Ohio State University; J.D., Capital University Law School; LL.M., University of Texas School of Law.

LLOYD T. WILSON, JR., *Associate Professor of Law.* B.A., Wabash College; M.A., Duke University; J.D., Indiana University School of Law—Bloomington.

MARY T. WOLF, *Clinical Professor of Law and Director of Clinical Programs.* B.A., Saint Xavier College; J.D., University of Iowa College of Law.

R. GEORGE WRIGHT, *Michael D. McCormick II Professor of Law.* A.B., University of Virginia; Ph.D., Indiana University; J.D., Indiana University School of Law—Indianapolis.

Emeriti

EDWARD P. ARCHER, *Professor of Law Emeritus.* B.M.E., Renesselaer Polytechnic Institute; J.D., LL.M., Georgetown University Law School.

JAMES F. BAILEY, III, *Professor of Law Emeritus.* A.B., J.D., M.A.L.S., The University of Michigan.

AGNES P. BARRETT, *Associate Professor of Law Emerita.* B.S., J.D., Indiana University.

CLYDE HARRISON CROCKETT, *Professor of Law Emeritus.* A.B., J.D., University of Texas; LL.M., University of London (The London School of Economics and Political Science).

DEBRA A. FALENDER, *Professor of Law Emerita.* A.B., Mount Holyoke College; J.D., Indiana University School of Law—Indianapolis.

DAVID A. FUNK, *Professor of Law Emeritus.* A.B., College of Wooster; J.D., Case Western Reserve University School of Law; M.A., The Ohio State University; LL.M., Case Western Reserve University; LL.M., Columbia Law School.

PAUL J. GALANTI, *Professor of Law Emeritus.* A.B., Bowdoin College; J.D., University of Chicago Law School.

HELEN P. GARFIELD, *Professor of Law Emerita.* B.S.J., Northwestern University; J.D., University of Colorado School of Law.

HAROLD GREENBERG, *Professor of Law Emeritus.* A.B., Temple University; J.D., University of Pennsylvania Law School.

WILLIAM F. HARVEY, *Carl M. Gray Professor of Law & Advocacy Emeritus.* A.B., University of Missouri; J.D., LL.M., Georgetown University Law School.

W. WILLIAM HODES, *Professor of Law Emeritus,* A.B., Harvard College; J.D., Rutgers University School of Law—Newark.

WILLIAM ANDREW KERR, *Professor of Law Emeritus.* A.B., West University; B.S., Duke University; J.D., LL.M., Harvard Law School.

WILLIAM E. MARSH, *Professor of Law Emeritus.* B.S., J.D., University of Nebraska.

RONALD W. POLSTON, *Professor of Law Emeritus.* B.S., Eastern Illinois University; LL.B., University of Illinois College of Law.

KENNETH M. STROUD, *Professor of Law Emeritus.* A.B., J.D., Indiana University—Bloomington.

JAMES PATRICK WHITE, *Professor of Law Emeritus.* A.B., University of Iowa; J.D., LL.M., George Washington University Law School.

Law Library Faculty

JUDITH FORD ANSPACH, *Professor of Law and Director, Ruth Lilly Law Library.* B.S., M.L.S., Kent State University; J.D., Mississippi College School of Law.

MARY HUDSON, *Reference/Circulation Librarian.* B.A., Ball State; M.L.S., Indiana University.

RICHARD HUMPHREY, *Reference Librarian.* A.A., Brewton-Parker Junior College; B.A., Georgia Southwestern College; M.L.S., University of Kentucky.

WENDELL E. JOHNTING, *Assistant Director for Technical Services.* A.B., Taylor University; M.L.S., Indiana University.

CHRIS E. LONG, *Catalog Librarian.* B.A., Indiana University; M.A., Indiana University; M.L.S., Indiana University.

MAHNAZ K. MOSHFEGH, *Acquisition/Serials Librarian.* B.A., National University of Iran; M.S., Tehran University; M.A., Ball State University; M.L.S., Ph.D., Indiana University.

MIRIAM A. MURPHY, *Associate Director of Law Library.* B.A., Purdue University; J.D., M.L.S., Indiana University—Bloomington.

KIYOSHI OTSU, *Computer System Specialist.* A.A., Parkland College; A.B., M.S., C.A.S., University of Illinois.

Indiana Law Review

Volume 37 **2003** **Number 1**

TABLE OF CONTENTS

ARTICLES

NOTES

Indiana Law Review

Volume 37 **2003** **Number 1**

ARTICLES

A TANGLED *WEBB*—REEXAMINING THE ROLE OF DUTY IN INDIANA NEGLIGENCE ACTIONS

HONORABLE THEODORE R. BOEHM*

INTRODUCTION

The four elements of a negligence action have long been recited by courts in Indiana and elsewhere as duty, breach, causation and harm. The American Law Institute's current consideration of the Restatement (Third) of Torts has produced a vigorous national debate over the roles of duty and causation in a negligence action.[1] Indiana's causation doctrine is relatively conventional, and although it presents its own set of problems, in general these are not unique to Indiana and are beyond the scope of this article except as duty and proximate cause relate to each other.[2] Duty, on the other hand, has a unique Indiana history that illuminates and in my view helps resolve the national debate. It alone is the subject of this article. In this article I offer my view on the reasons courts often inappropriately speak in terms of duty when the issue really is something else. I believe negligence actions are best understood by recognizing that every actor has an obligation to behave reasonably. This approach is not designed either to expand or to constrict liability. Rather it focuses on what I believe to be the central legal issue in a negligence claim: are there any factors—usually dubbed policy considerations—that preclude this claimant as a matter of law from recovering

---

* Justice, Indiana Supreme Court; A.B., 1960, Brown University; J.D., 1963, Harvard Law School. I would like to thank my law clerks, Michael Limrick, Allison Brown, Mildred Van Volkom, and Paul Jefferson, and my assistant Debra Moss, for their help in preparing this article.

1. *See generally Symposium: The John W. Wade Conference on the Third Restatement of Torts*, 54 VAND. L. REV. 3 (2001).

2. Proximate cause is equally subject to criticism, but for purposes of this article I restrict my comments to the court's role in allowing or denying a claim under the duty rubric. I use "causation" rather than "scope of liability" not because I prefer that terminology, but only because it is conventional and avoids distractingly long diversions from the points I seek to make. It seems widely understood that "proximate cause" embraces both causation in fact and a legal concept variously described as "legal cause," foreseeability, scope of liability, etc. Although I agree with those who would abolish the term as confusing to the jury and to analysis by appellate courts, I nevertheless use it as shorthand. The alternative is to fill every discussion of scope of liability with an explication of precisely how the term is used, even though it is irrelevant to the analysis of the court's role in assessing the presence or absence of "duty."

from this defendant under these circumstances?

I. DUTY AS AN ELEMENT OF A NEGLIGENCE CLAIM

The concept of duty has become increasingly subject to criticism in general tort literature as either wholly unnecessary or hopelessly confused. Although some view duty as a concept that is "perfectly intuitive" and "central to negligence law,"[3] others find its application "so changeable that it actually defies the idea of a definition."[4] Indiana's approach to duty has been no less problematic. The charge has been leveled that, over the years, Indiana courts have found duties to be present or absent "independently and haphazardly, without any thought given to their relationship to other tort obligations arising in other factual contexts."[5]

A few things seem clear. "Duty" is used to describe a rule set down by the courts and applied by the judge to permit a claim to go forward. In lawyer's terms, it is a question of law. If the court finds no duty of the defendant to the plaintiff, that is the end of the plaintiff's negligence claim. Lack of "causation" or "proximate cause" may also serve to deny a claim, but that determination is ordinarily for the trier of fact. It is also clear that many specific duties have been found by the courts to be owed by landowners, occupiers of land and social hosts,[6] but those same parties have been said in other circumstances to owe no duty while engaged in the same or similar activity.[7] The same is true for, among others, health care providers,[8] utilities,[9] railroads,[10] motorists,[11] automobile

---

3. John C.P. Goldberg, *Duty & the Structure of Negligence*, 10 KAN. J.L. & PUB. POL'Y 149, 150 (2000).

4. Peter F. Lake, *Common Law "Duty" Analysis: The Conceptual Expansion of "Duty" in a Period of Doctrinal Consolidation/Retrenchment*, 10 KAN. J.L. & PUB. POL'Y 153, 154 (2000).

5. Jay Tidmarsh, *Tort Law: The Languages of Duty*, 25 IND. L. REV. 1419, 1425 (1992).

6. Delta Tau Delta v. Johnson, 712 N.E.2d 968, 971-73 (Ind. 1999) (social host owes duty to take reasonable precautions to protect invitees from foreseeable criminal attacks); Tibbs v. Huber, Hunt & Nichols, Inc., 668 N.E.2d 248, 250 (Ind. 1996) (recognizing that a contractor owes a duty to passersby to keep adjoining areas reasonably clear of risks); Valinet v. Eskew, 574 N.E.2d 283, 285-86 (Ind. 1991) (finding possessor of land in urban area owes duty to prevent unreasonable risk of harm arising from condition of trees near the highway); Lutheran Hosp. of Ind., Inc. v. Blaser, 634 N.E.2d 864, 868-69 (Ind. Ct. App. 1994) (holding a landowner owes a duty to exercise reasonable care to discover defects or dangerous conditions on the premises, and invitor owes duty to provide safe means of ingress and egress); Gunter v. Vill. Pub, 606 N.E.2d 1310, 1312 (Ind. Ct. App. 1993) (finding that a tavern owner owes a duty to protect patrons from reasonably foreseeable disorderly acts of other patrons).

7. Van Duyn v. Cook-Teague P'ship, 694 N.E.2d 779, 781 (Ind. Ct. App. 1998) (recognizing that generally there is no duty on part of business owners to protect patrons from criminal acts of third persons); Fawley v. Martin's Supermarkets, Inc., 618 N.E.2d 10, 14 (Ind. Ct. App. 1993) (requiring no duty to invitees to protect from runaway vehicles crossing onto pedestrian sidewalks from parking lots).

8. *Compare* Harris v. Raymond, 715 N.E.2d 388, 394-95 (Ind. 1999) (holding that

passengers,[12] schools,[13] employers and contractors,[14] insurers,[15] the government,[16]

---

physicians have duty to warn current and former patients of safety issues highlighted either by manufacturer of medical device or the FDA), *and* Walker v. Rinck, 604 N.E.2d 591, 594-95 (Ind. 1992) (holding that physician owes duty to later-born children of mother with Rh-negative blood to provide the mother RhoGAM following birth of Rh-positive child), *and* Cowe v. Forum Group, Inc., 575 N.E.2d 630, 636-37 (Ind. 1991) (holding that nursing home owes total care patients duty to exercise reasonable care for their protection, and owes patient's unborn children duty because of extreme dependence of patient upon nursing home for that care), *with* Auler v. Van Natta, 686 N.E.2d 172, 175 (Ind. Ct. App. 1997) (recognizing absent circumstances supporting claim for vicarious liability or other special circumstances, hospital has no independent duty to obtain patient's informed consent).

9. *Compare* Bush v. N. Ind. Pub. Serv. Co., 685 N.E.2d 174, 177 (Ind. Ct. App. 1997) (finding a duty to those persons using road as it was intended to be used), *and* Goldsberry v. Grubbs, 672 N.E.2d 475, 480 (Ind. Ct. App. 1996) (recognizing a duty to motoring public to exercise reasonable care when placing poles along highways), *with* Butler v. City of Peru, 733 N.E.2d 912, 916 (Ind. 2000) (holding a utility company generally owes no duty to those injured by power lines owned by its customers).

10. *Compare* CSX Transp., Inc. v. Kirby, 687 N.E.2d 611, 615 (Ind. Ct. App. 1997) (recognizing there is a duty to construct and maintain crossings so that they will be reasonably safe for travel), *with* Cent. Ind. Ry. Co. v. Anderson Banking Co., 240 N.E.2d 840, 849 (Ind. App. 1968) (holding there is no duty to equip crossings with automatic signals or provide reflectors on sides of cars unless required to do so by Public Service Commission).

11. *Compare* Stephenson v. Ledbetter, 596 N.E.2d 1369, 1372 (Ind. 1992) (requiring a duty to passengers to exercise reasonable care in automobile's operation), *with* Merida v. Cardinal, 749 N.E.2d 605, 607 (Ind. Ct. App. 2001) (holding there is no duty of a driver on "preferred" street to look left and right at intersection when no notice other driver would violate law).

12. *Compare Stephenson*, 596 N.E.2d at 1372 (recognizing a duty to use reasonable care to avoid injury to self), *with* Hopper v. Carey, 716 N.E.2d 566, 574 (Ind. Ct. App. 1999) (holding there is no duty to wear seatbelt).

13. *Compare* Norman v. Turkey Run Cmty. Sch. Corp., 411 N.E.2d 614, 616 (Ind. 1980) (imposing duty to supervise students), *with id.* at 618 (finding no duty to pay particular attention to particular student running on playground).

14. *Compare* Kostidis v. Gen. Cinema Corp., 754 N.E.2d 563, 567 (Ind. Ct. App. 2001) (recognizing a construction contract implies a duty to do work skillfully, carefully, and in a workmanlike manner), *with* Blake v. Calumet Constr. Corp., 674 N.E.2d 167, 170 (Ind. 1996) (holding contractors owe no duty to third parties after owner has accepted work, except where work is deemed dangerously defective, inherently dangerous or imminently dangerous).

15. *Compare* Stump v. Commercial Union, 601 N.E.2d 327, 332 (Ind. 1992) (finding a duty not to handle claims in grossly negligent manner), *with* Baxter v. Ind. State Teacher's Ass'n Ins. Trust, 749 N.E.2d 47, 57 (Ind. Ct. App. 2001) (holding there is no duty to protect second insurer's subrogation rights).

16. *Compare* Benton v. City of Oakland City, 721 N.E.2d 224, 233 (Ind. 1999) (requiring a duty to maintain public recreational facility in reasonably safe manner), *with* Weatherholt v. Spencer County, 639 N.E.2d 354, 356 (Ind. Ct. App. 1994) (finding the county owed no duty to inmate to administer its jail).

polygraph examiners,[17] and animal keepers.[18] The common thread, of course, is that the presence of a "duty" determines whether the plaintiff may recover from the defendant, assuming causation and breach are proved. But the approach to finding duty varies widely. Duty may be found or rejected based on who the plaintiff is, the standard of conduct required of the defendant, or the circumstances under which the defendant acted. For example, a trespasser may not sue a landowner for a negligently maintained property but an invitee may.[19] A client may assert a claim against a lawyer for negligence, but a non-client may not.[20] A social host has no common law duty to refuse to serve a potentially inebriated guest, but a bartender does.[21] Duties in these examples have been rejected or embraced on qualitatively different considerations, including concern for an unmanageable scale of exposure, distaste for the plaintiff, and perceived social norms. The result of these rules addressing various fact situations is a series of data points that adhere to no common logic: a mosaic that forms no pattern.

Although it is less objectively demonstrable, I believe one can find many instances in Indiana case law where the case is resolved by finding no duty, but it would seem more accurate to identify the reason the plaintiff lost as a failure to establish either unreasonable conduct or causation. As a single example, take *Norman v. Turkey Run Community School Corp.*,[22] where the Indiana Supreme Court held that teachers owed no duty to a four-year-old who was injured when he ran into another child at a playground under the teachers' supervision. A more accurate reason for dismissing the plaintiff's case would seem to be that the teachers were not negligent as a matter of law because the particular injury simply was not preventable by any reasonable precautions. I suggest Indiana cases offer many examples of the same phenomenon—finding no duty where some less sweeping fact-specific reason would better explain the plaintiff's failure to carry the day.

---

17. Lawson v. Howmet Aluminum Corp., 449 N.E.2d 1172, 1177 (Ind. Ct. App. 1983) (holding there is a duty to examinee to exercise reasonable care in conducting polygraph examination).

18. Blake v. Dunn Farms, Inc., 413 N.E.2d 560, 563 (Ind. 1980) (requiring a duty to provide for restraining and confinement of animal).

19. Frye v. Trs. of Rumbletown Free Methodist Church, 657 N.E.2d 745, 748 (Ind. Ct. App. 1995).

20. *Compare* Rice v. Strunk, 670 N.E.2d 1280, 1283-84 (Ind. 1996) (finding duty to exercise ordinary skill and knowledge), *with* Hacker v. Holland, 570 N.E.2d 951, 955 (Ind. Ct. App. 1991) (holding generally there is no duty to non-clients).

21. *Compare* Gariup Constr. Co., Inc. v. Foster, 519 N.E.2d 1224, 1228 (Ind. 1988) (requiring no duty of social host), *with* Elder v. Fisher, 217 N.E.2d 847, 853 (1966) (recognizing common law duty of sellers of alcohol). Apart from these common law rules, liability may or may not also be established for damages resulting from providing an inebriated person alcohol under section 7.1-5-10-15.5 of the Indiana Code. IND. CODE § 7.1-5-10-15.5 (1998).

22. 411 N.E.2d 614 (Ind. 1980).

II. *WEBB V. JARVIS* AND ITS SUBSEQUENT APPLICATION

In 1991, in *Webb v. Jarvis*,[23] the Indiana Supreme Court, in an attempt to establish a consistent formula to identify a duty, announced a tripartite test that directed trial courts to balance (1) the relationship between the parties, (2) the foreseeability of harm, and (3) public policy concerns. But preexisting notions of duty were not so easily eradicated. Since *Webb*, courts have concluded that its formula supersedes those individual strains of duty, that it complements them, or that it applies only when a new issue of duty arises. In many cases, *Webb* has been ignored; in others, it has been misapplied.

Webb involved a patient who had been over-prescribed steroids and in a rage shot his brother-in-law. The brother-in-law sued the prescribing doctor, arguing that the doctor breached a duty to administer medical treatment in such a way as to take into account possible harm to others. The Indiana Supreme Court held that the doctor owed no such duty.[24] The court concluded that all three factors under the newly formulated test weighed against the imposition of a duty. At the time, the *Webb* balancing test was "new" in the sense that "Indiana cases had long recognized the need of a tort plaintiff to establish a duty, but no single test to determine the existence of a duty had ever been established."[25]

A. Early Inconsistent Application of Webb

From the beginning there was uncertainty as to whether *Webb* supplied the only test to determine whether a duty exists or whether the long-established common law rules of duty also survived. As one commentator soon noted, "[i]t is easy to declare, as *Webb v. Jarvis* does, a new test for duty. It is difficult to apply that analysis to existing duty rules, many of which cannot be justified under *Webb*'s analysis."[26] In fact, the confusion over *Webb*'s proper role began with an opinion the Indiana Supreme Court had issued only three days earlier in *Valinet v. Eskew*.[27] *Valinet* involved a claim by a passing motorist hit by a falling limb, and addressed the question of a landowner's duty to maintain the trees on the property. Although *Webb* and *Valinet* were decided nearly simultaneously, the court in *Valinet* simply adopted the rule of section 363 of the Restatement (Second) of Torts that possessors of land in rural areas are not liable for physical harm to others resulting from the condition of trees near a highway.[28] There was no discussion of the relationship, foreseeability, and public policy factors, or their possible effect on the outcome of the case.

Through the years, the most consistent criticism of *Webb* has been the contention that the issue of duty cannot fit neatly within a three-part balancing

---

23. 575 N.E.2d 992 (Ind. 1991).
24. *Id.* at 995.
25. Tidmarsh, *supra* note 5, at 1424-25 (footnote omitted).
26. *Id.* at 1466-67.
27. 574 N.E.2d 283 (Ind. 1991).
28. RESTATEMENT (SECOND) OF TORTS § 363 (1965).

test. That assertion has its Hoosier roots in *Gariup Construction Co. v. Foster*,[29] which predated *Webb* by three years. In *Gariup*, the Indiana Supreme Court quoted, with approval, the following passage from *Prosser & Keeton on the Law of Torts*:

> It is . . . not surprising to find that the problem of duty is as broad as the whole law of negligence, and that no universal test for it ever has been formulated. . . . But it should be recognized that "duty" is not sacrosanct in itself, but is only an expression of the sum total of those considerations of policy which lead the law to say that the plaintiff is entitled to protection. . . . No better general statement can be made than that the courts will find a duty where, in general, reasonable persons would recognize it and agree that it exists.[30]

This formulation—duty is no more than the sum of the policy considerations bearing on the plaintiff's right to recover—is, I think, still valid. It is another way to say that duty, or the absence of duty, is an expression of the result of the analysis, not a tool used to reach that result.

Despite *Webb*'s attempt to establish a standard test for duty, one month later, in *Cowe v. Forum Group, Inc.*,[31] the same court examined whether a medical care provider owed a duty to an unborn fetus to detect its mother's pregnancy. Instead of employing the *Webb* analysis, the court found a duty analogous to "that of a common carrier to provide protection and care."[32] That duty was owed to the mother and, by extension, to her unborn child.[33] The court continued to employ traditional duty rules the next year in *Stephenson v. Ledbetter*,[34] in which it held that the driver of a pickup truck owed a duty of reasonable care to an intoxicated passenger who fell out of the pickup's bed.[35] Although the court cited both *Webb* and *Gariup* for the proposition that duty is a question of law, it referred to neither case when it came time to determine whether a duty existed. Instead, the court relied on "the common law view . . . that the operator of an automobile owes to a passenger the duty of exercising reasonable care in its operation."[36] That same year, in *Stump v. Commercial Union*,[37] the court treated *Webb* as just one way to determine the existence of a duty, not as an exclusive test, and then applied the *Webb* factors to determine that a workers' compensation insurance carrier owed a duty to injured employees not to handle claims in a grossly negligent manner.[38]

---

29. 519 N.E.2d 1224 (Ind. 1988).

30. *Id.* at 1227 (quoting W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 53, at 357-59 (5th ed. 1984)).

31. 575 N.E.2d 630 (Ind. 1991).

32. *Id.* at 636.

33. *Id.* at 637.

34. 596 N.E.2d 1369 (Ind. 1992).

35. *Id.* at 1372.

36. *Id.* (citing Munson v. Rupker, 148 N.E. 169 (1925)).

37. 601 N.E.2d 327 (Ind. 1992).

38. At the same time, the Indiana Court of Appeals also demonstrated a reluctance to employ

In 1996, the Indiana Supreme Court decided three cases that seemed to suggest a drift away from *Webb*. In *Tibbs v. Huber, Hunt & Nichols, Inc.*,[39] a state employee, injured when he slipped on a pipe in a stairway, sued the general contractor and subcontractor cutting pipe in a nearby hallway. The subcontractor claimed no duty because it did not control the stairwell. The court hinged its analysis on the foreseeability of injury, instead of balancing the *Webb* factors: "[T]he stairwell was certainly within the 'range of apprehension' and, accordingly, [the defendant] was obliged to behave safely."[40] This seems more properly an exercise in scope of liability than a duty analysis. It assumes the contractor and its subcontractor could be liable to the employee as a person who could be expected to use the stairway and therefore seems to assume a duty on the part of the contractor and subcontractor to maintain the stairway in a safe condition.

Similarly, in *Rice v. Strunk*,[41] the court assumed that the "duty" imposed on an attorney sued for negligence required an attorney-client relationship and rejected balancing the relationship of the parties against the other two *Webb* factors:

> [B]ecause it is necessary . . . for the plaintiffs to show the existence of an attorney-client relationship, the existence of duty will turn initially on the relationship of the parties. That is, if an attorney-client relationship does not exist, it will not be necessary to reach the foreseeability and public policy factors.[42]

Finally, in *Blake v. Calumet Construction Corp.*,[43] the court acknowledged that it "usually considers three factors"[44] when determining duty, but the court did not describe *Webb* as the exclusive test. Instead, it noted William Prosser's article, *Palsgraf Revisited*,[45] which discussed "other considerations, such as conscience of the community and ease of administration."[46] In the end, the court noted that the application of a century-old common law rule was not challenged by either party and for that reason affirmed the trial court under a traditional

---

Webb as the exclusive test for duty. *See, e.g.*, Gunter v. Vill. Pub, 606 N.E.2d 1310, 1312 (Ind. Ct. App. 1993) (citing *Burrell v. Meads*, 569 N.E.2d 637 (Ind. 1991), for the proposition that "[i]n Indiana, landowners have a duty to exercise reasonable care to make their premises safe for business invitees," and *Welch v. Railroad Crossing, Inc.*, 488 N.E.2d 383, 388 (Ind. Ct. App. 1986), for the proposition that "[a] duty to anticipate and to take steps to protect against a criminal act arises only when the *facts of a particular case* make it reasonably foreseeable that a criminal act is likely to occur.").

39. 668 N.E.2d 248 (Ind. 1996).
40. *Id.* at 250.
41. 670 N.E.2d 1280 (Ind. 1996).
42. *Id.* at 1284.
43. 674 N.E.2d 167 (Ind. 1996).
44. *Id.* at 170 (quoting Webb v. Jarvis, 575 N.E.2d 992, 995 (Ind. 1991)).
45. 52 MICH. L. REV. 1 (1953).
46. *Blake*, 674 N.E.2d at 170.

common law principle: "The duty inquiry in this case . . . is governed by a line of decisions dealing specifically with contractors' liability to third parties for construction flaws."[47]

In *Walker v. Rinck*,[48] the Indiana Supreme Court applied the *Webb* formulation in a pre-conception medical malpractice case. In *Walker*, later-born children sued a doctor for failing to give RhoGAM injections to their mother, who had Rh-negative blood, after the birth of her first child. The court determined: (1) the children "were the beneficiaries of the consensual relationship" between their mother and her doctor; (2) the various health deficiencies suffered by the children were foreseeable; and (3) the public policy behind the issuance of RhoGAM injections is "to protect future children."[49] Therefore, a duty existed.

In *Erie Insurance Co. v. Hickman*,[50] the court applied *Webb* not to determine the existence of a duty, but rather to determine whether Indiana recognized a tort cause of action based on breach of an already recognized contractual duty. The court first observed that "Indiana law has long recognized that there is a legal duty implied in all insurance contracts that the insurer deal in good faith with its insured."[51] In determining whether the breach of that duty constituted a tort, the court turned to *Webb*. Citing the "unique character" of the insurance relationship, the foreseeability of harm to an insured, and the public policy of "fair play between insurer and insured," the court held that "recognition of a cause of action for the tortious breach of an insurer's duty to deal with its insured in good faith is appropriate."[52]

B. Webb *in the Court of Appeals*

The Indiana Court of Appeals has generally, but not uniformly, adhered to *Webb* as the test for duty.[53] Notably, in 1994, the court of appeals per then Judge

---

47. *Id.* at 170.

48. 604 N.E.2d 591 (Ind. 1992).

49. *Id.* at 594-95.

50. 622 N.E.2d 515 (Ind. 1993).

51. *Id.* at 518 (citing Vernon Fire & Cas. Ins. Co. v. Sharp, 349 N.E.2d 173, 181 (Ind. 1976)).

52. *Id.* at 518-19.

53. *See, e.g.*, Zurich Am. Ins. Group v. Wynkoop, 746 N.E.2d 985, 990 (Ind. Ct. App. 2001) (stating that "[t]his court has analyzed the question of what must be considered in order for a court to impose a duty at common law. Three factors must be balanced . . . ."); Ousley v. Bd. of Comm'rs, 734 N.E.2d 290, 294 (Ind. Ct. App. 2000) (holding that "the *Webb* test has become firmly entrenched in the case law of Indiana."); Ind. Bell Tel. Co. v. Maynard, 705 N.E.2d 513, 514 (Ind. Ct. App. 1999) (stating "[i]n determining whether to impose a duty in a negligence action at common law, this court considers three factors . . . ."); Ebbinghouse v. FirstFleet, Inc., 693 N.E.2d 644, 647 (Ind. Ct. App. 1998) (citing *Webb*, 575 N.E.2d at 995) (reasoning "[o]ur supreme court has held that in determining whether a defendant owed a duty to the plaintiff, we must consider and balance three factors . . . ."); *but see* Basicker *ex rel.* Johnson v. Denny's, Inc., 704 N.E.2d 1077, 1080 (Ind. Ct. App. 1999) (citing Van Duyn v. Cook-Teage P'ship, 694 N.E.2d 779, 781 (Ind. Ct.

Rucker, applied the *Webb* formulation in examining whether counselors and clergy with knowledge of a child's sexual abuse owed a duty to the child to report the abuse.[54] The court held that: (1) no significant relationship arose from either the knowledge of abuse, or the defendant's status as a marriage counselor of the abuser or clergy for the victim; (2) it was foreseeable that the abuse, if unreported, would continue; and (3) public policy was reflected in legislation[55] that criminalized failure to report but had created no civil cause of action.[56] Balancing the factors, the court concluded that no duty existed on the part of three of four defendants to report the abuse. That same year, the Indiana Supreme Court issued two opinions that seemed to come down on the side of *Webb* as the exclusive test for duty. The court applied the *Webb* test to determine the existence of a "private duty of a governmental entity,"[57] and later that year, again applied *Webb* in holding that pharmacists owe a duty of reasonable care to customers to stop filling prescriptions when they know, or should know from the frequency of prescription orders, that the drugs are being misused.[58]

Although attempting to apply *Webb* as a test for duty, the court of appeals has taken a variety of views as to how this is to be done. The efforts of the court of appeals to rationalize and apply *Webb* are discussed in Part III.A.

C. Recent Explanations of Webb

The Indiana Supreme Court has continued to waffle over the proper application of *Webb*. In recent years, it has described *Webb* as a "useful," though

---

App. 1998)) (relying on traditional duty concept that "[g]enerally, there is no duty on the part of a business owner to protect its patrons against the criminal acts of third persons unless the particular facts make it reasonably foreseeable that the criminal act will occur"); CSX Transp., Inc. v. Kirby, 687 N.E.2d 611, 615 (Ind. Ct. App. 1997) (citing Penn. R.R. v. Mink, 212 N.E.2d 784, 788 (Ind. App. 1966)) (relying on traditional rule that "[r]ailroad companies owe to the traveling public the duty of exercising reasonable care to avoid injury to persons at places where the tracks and the highway cross"); Sheley v. Cross, 680 N.E.2d 10, 12 (Ind. Ct. App. 1997) (relying on traditional rule that "a landowner does owe a duty to the traveling public to exercise reasonable care in the use of his property so as not to interfere with safe travel on public roadways").

54. J.A.W. v. Roberts, 627 N.E.2d 802 (Ind. Ct. App. 1994).

55. IND. CODE § 31-6-11-20 (1993) (recodified at IND. CODE § 31-33-22-1 (1998)).

56. *Roberts*, 627 N.E.2d at 813.

57. Mullin v. Mun. City of South Bend, 639 N.E.2d 278, 283 (Ind. 1994). That application prompted a dissent from Justice Dickson, who reiterated his position from *Gariup Construction Co.*:

> The legal determination of whether a duty exists in a particular case is not necessarily resolved by recourse to the three factors found useful by this Court in *Webb v. Jarvis*. These should not constitute the exclusive tests for duty. . . . The majority opinion presents an unnecessary and unwise construct. It is better that the common law avoid such artificial and rigid formulations.

Id. at 285-86 (Dickson, J., concurring and dissenting) (citations omitted).

58. Hooks Super X, Inc. v. McLaughlin, 642 N.E.2d 514 (Ind. 1994).

"not exclusive," test for duty,[59] as the only test,[60] and as the "usual" test.[61]

In *Delta Tau Delta v. Johnson*,[62] the court addressed the duty of a college fraternity to protect invitees from the criminal acts of third parties. The court first stated that it "need not formally use the three factor balancing test as enunciated in *Webb v. Jarvis*" because the fraternity already had a duty as a landowner "to exercise reasonable care for [the plaintiff's] protection."[63] However, the court then took the additional step of stating that "[t]he issue in this case is when, if ever, does that duty extend to criminal acts by third parties."[64] For this question, the court returned to the *Webb* framework. In that context, the court ultimately adopted a "totality of the circumstances" test to determine whether the injury to the plaintiff—a rape by another guest—was foreseeable.[65] In one respect, *Delta Tau Delta* foreshadowed *NIPSCO v. Sharp*,[66] where the majority of the Indiana Supreme Court explicitly disapproved the *Webb* analysis used by the court of appeals to identify the duty of a supplier of electrical power. The supreme court found the duty to keep power lines insulated where the public may come into contact with them was found to had been long ago settled by precedent. The court then explained that where a "duty has already been declared" resort to *Webb* is unnecessary.[67]

Shortly before *NIPSCO* was decided, the Indiana Supreme Court, in *Mangold ex rel. Mangold v. Indiana Department of Natural Resources*,[68] quoted a passage from *Webb* dealing with its third "public policy" factor. *Webb* in turn quoted Prosser and Keeton describing duty as "only an expression . . . of those considerations of policy which lead the law to say that the plaintiff is entitled to protection."[69] In my view, the Prosser and Keeton quotation supports the broader statement that policy considerations are the only relevant factor necessary in determining duty, not just one of three separate and independent factors.

Although reiterating the *Webb* formula, the court in *Mangold* moved closer to the view that duty is merely the label a court pins on a defendant as the result of an evaluation of policy considerations. The court expressly recognized that "[b]y declaring that a school may be held liable for the injuries suffered by its students, we essentially have made a policy decision that a school's relationship to its students, the foreseeability of harm, and public policy concerns entitle

---

59. Cram v. Howell, 680 N.E.2d 1096, 1097 n.1 (Ind. 1997).

60. *Rice*, 670 N.E.2d at 1284 n.1.

61. *Blake*, 674 N.E.2d at 170.

62. 712 N.E.2d 968 (Ind. 1999).

63. *Id.* at 971 n.4.

64. *Id.*

65. *Id.* at 973.

66. 790 N.E.2d 462 (Ind. 2003). The court of appeals addressing the issue in *Webb* terms is found at *NIPSCO v. Sharp*, 732 N.E.2d 848, 856-58 (Ind. Ct. App. 2000).

67. *NIPSCO*, 790 N.E.2d at 465.

68. 756 N.E.2d 970 (Ind. 2001).

69. *Webb*, 575 N.E.2d at 997.

students to protection."[70] This formulation recognizes that foreseeability of harm and the relationship between the parties are factors that may bear on the ultimate "policy decision," whether the law allows recovery under these circumstances. The *Mangold* opinion pointed out that although the existence of duty is a question of law, breach of the duty is ordinarily a question for the trier of fact.[71] In this respect, the court cited *Bader v. Johnson*,[72] which reformulated the elements of a negligence claim from: (1) duty, (2) breach, (3) causation, and (4) damages to (1) duty, (2) breach, and (3) "compensable injury proximately caused" by the breach.[73] This formulation requires that damages be "compensable" and be "proximately caused." It thus arguably presages an explicit compression of proximate cause and damages into one requirement that the harm be within the "scope of liability" created by the tortious act.

III. THE PROBLEM WITH *WEBB*

Webb and its progeny do not attempt to explain how the three factors interact and how they are to be balanced. But the principal difficulty in applying the *Webb* formula lies in the factors it purports to "balance." Public policy has always been the linchpin of duty, and remains so under the formulation I suggest. The other two factors—relationship and foreseeability—are either subsumed into the policy issue or more properly viewed as bearing on issues for the trier of fact. To be sure, the relationship between the parties may bear on whether the law will allow the plaintiff to collect from the defendant—consider the trespasser suing the landowner—but that relationship is not a distinct factor as opposed to one of many considerations that may affect the policy call.

A. Foreseeability as a Webb *Factor*

Foreseeability, on the other hand, is problematic for a number of reasons. First, there is no agreement as to what that term means under *Webb*. In 1996, in *Goldsberry v. Grubbs*,[74] Judge Kirsch suggested that foreseeability in the duty context was distinct from foreseeability for purposes of proximate cause. The court in *Goldsberry* reasoned that because foreseeability is applied in both duty and proximate cause it must mean something different in the two contexts.[75] If it were not two distinct concepts, as Judge Bailey recently put it, "deciding the duty question would subsume the entire law of negligence, i.e., duty breach and proximate cause, into the duty question."[76] To resolve this difficulty, Judge Kirsch explained that in a duty analysis under *Webb*, foreseeability is viewed

---

70. *Mangold*, 756 N.E.2d at 974.

71. *Id.* at 975.

72. 732 N.E.2d 1212 (Ind. 2000).

73. *Id.* at 1216-17.

74. 672 N.E.2d 475 (Ind. Ct. App. 1996).

75. *Id.* at 479.

76. Hammock v. Red Gold, Inc., 784 N.E.2d 495, 507 (Ind. Ct. App. 2003) (Bailey, J., dissenting) (quoting *Goldsberry*, 672 N.E.2d at 479).

prospectively from the perspective of the actor and without regard to the facts of the particular case. In contrast, foreseeability for purposes of proximate cause is viewed retrospectively, and takes into account the particular circumstances of each case.[77] Judge Friedlander dissented in *Goldsberry* and explicitly rejected the suggestion that the label "foreseeability" was pinned on different concepts in the two contexts.[78] In Judge Friedlander's view, imposition of a duty necessarily embraced the specific circumstances of the actor, and the court should adhere to the classic formulation of duty to exercise reasonable care as owed only to those reasonably foreseeably injured by a breach.[79]

Since 1996, the schizophrenic view of foreseeability proposed by the *Goldsberry* majority has been embraced by some panels of the Indiana Court of Appeals,[80] rejected by others,[81] and noted by some without taking sides.[82] *Goldsberry* seems accurately described by Judge Bailey as an effort to address "some of the confusion created by the Webb decision . . . ."[83] Attempting to straighten things out is the most the court of appeals can do when faced with directly applicable Indiana Supreme Court precedent. The logical result of a single "foreseeability" was to collapse proximate cause into duty. Because *Webb* distinguished the two, the court of appeals felt compelled to attempt to explain how the two can coexist. As explained in Part IV, I suggest that the logic of this situation drives us not to find two concepts of foreseeabilty, but rather to recognize that duty adds nothing to the analysis of a negligence action.

Moreover, I do not believe the effort to rationalize *Webb* by bifurcating foreseeability is successful. Like "proximate cause," "foreseeability" as a concept in tort law also has its critics. That, like the issues surrounding "proximate cause," is a subject for another day. Notwithstanding any shortcomings, "foreseeability" remains in common use and is specifically demanded by *Webb*. I think there are two problems in the two-foreseeabilities approach to *Webb*. First, I suspect it is not always easy and sometimes impossible to distinguish general circumstances from specific facts of a given case. Assuming that can be done, "prospective" general foreseeability is not restricted to duty issues.

I take prospective foreseeability to mean a reasonable person in the

---

77. *Goldsberry*, 672 N.E.2d at 479.

78. *Id.* at 483 (Friedlander, J., dissenting).

79. *Id.* at 482 (Friedlander, J., dissenting) (citing NIPSCO v. Sell, 597 N.E.2d 329, 332 (Ind. Ct. App. 1992) (where the court of appeals, citing Justice Cardozo's classic *Palsgraf* analysis, adopted this formulation of duty under *Webb*)).

80. City of Gary v. Smith & Wesson, 776 N.E.2d 368, 386 (Ind. Ct. App. 2002); City of Indianapolis v. Pippin, 726 N.E.2d 341, 346 (Ind. Ct. App. 2000); Franklin v. Benock, 722 N.E.2d 874, 879 (Ind. Ct. App. 2000); King v. Northeast Sec., Inc., 732 N.E.2d 824, 834 (Ind. Ct. App. 2000).

81. *Hammock*, 784 N.E.2d at 501 n.10; Bradtmiller v. Hughes Prop., Inc., 693 N.E.2d 85, 89 (Ind. Ct. App. 1998); Bush v. NIPSCO, 685 N.E.2d 174, 179 (Ind. Ct. App. 1997) (dissenting Judge Rucker expressly agreeing with *Goldsberry*).

82. Ousley v. Bd. of Comm'rs of Fulton County, 734 N.E.2d 290 (Ind. Ct. App. 2000).

83. *Hammock*, 784 N.E.2d at 506 (Bailey, J., dissenting).

defendant's shoes who thought about it would recognize a reasonable possibility that harm of the sort the plaintiff suffered would result from the defendant's act or omission. So understood, "foreseeability" may well bear on whether the plaintiff's injury is among the risks created by the defendant's unreasonable behavior for which the law should provide relief. That is a component of proximate cause as currently understood. As the comments to the pattern jury instruction observe, "the trier of fact considers if the injury was a natural and probable consequence of a negligent act, which in the light of the circumstances could have been reasonably foreseen."[84] Causation in fact is also embraced within proximate cause as typically instructed in Indiana courts and elsewhere. Indeed, the pattern instruction language focuses on causation in fact, not scope of liability.[85] Determination of causation in fact is a retrospective analysis. But the scope of liability component of proximate cause does not turn on a retrospective view of foreseeability, and it may also take into consideration the specific circumstances under which the defendant acted. This turns on whether the harm incurred by the plaintiff is within the risks that make the defendant's conduct unreasonable. It nevertheless is a component of proximate cause and is an issue for the trier of fact.[86] Finally, foreseeability of harm may also relate to whether the defendant's conduct fell below the reasonableness standard. This is a breach issue in conventional duty/breach analysis. Neither of these uses of foreseeability makes it an independent factor in whether the defendant's action as a matter of law is or is not a basis for a claim. Both may invoke specific circumstances, and evaluate the conduct from the perspective of the actor.

B. Duty, Scope of Liability, and Breach

Webb's reference to foreseeability often causes confusion between duty and breach as well as between duty and proximate cause. Foreseeability is described by *Webb* as an independent factor in determining duty, but it may also relate to whether the defendant's actions were reasonable, i.e. to the defendant's breach. For example, a shopkeeper whose business is in a low-crime area and whose patrons have never been attacked may argue that his or her decision not to hire security guards was reasonable given the circumstances.[87] More frequently, foreseeability is also a critical component of causation. Thus, courts have found that a defendant owed no duty to a plaintiff when, as a matter of law, the

---

84. IND. PATTERN JURY INST. CIVIL cmts. (2d ed.) (citing Hedrick v. Tabbert, 722 N.E.2d 1296 (Ind. Ct. App. 1999); Mangold v. Ind. Dept. of Natural Resources, 720 N.E.2d 424 (Ind. Ct. App. 1999)).

85. Section 5.06 provides: "'Proximate cause' is that cause which produces the [death][injury][property damage] complained of and without which the result would not have occurred. That cause must lead in a natural and continuous sequence to the resulting [death][injury][property damage][unbroken by any intervening cause]." *Id.* § 5.06.

86. *See* Rubin v. Johnson, 550 N.E.2d 324, 331 (Ind. Ct. App. 1990).

87. *Cf.* Basicker *ex rel.* Johnson v. Denny's, Inc., 704 N.E.2d 1077, 1080 (Ind. Ct. App. 1999).

defendant was not the proximate cause of the plaintiff's injury. That was at least part of the issue in *Webb*, which is itself an illustration of the conflation of these factors. *Webb* characterized its analysis as focusing "on whether the person actually harmed was a foreseeable victim and whether the type of harm actually inflicted was reasonably foreseeable."[88] Although we may argue over whether "foreseeable" is the best way to describe harms within the risks created by the defendant's act, the formulation in *Webb* is a fair description of the inquiry that is left to the jury in determining whether the defendant acted unreasonably and, if so, whether that unreasonable act was the proximate cause of injury. As the Indiana Supreme Court recently put it:

> Under Indiana law, a negligent defendant may be liable for a plaintiff's injury if his or her action is deemed to be a proximate cause of that injury. Whether or not proximate cause exists is primarily a question of foreseeability. As this Court recently stated, the issue is whether the injury "is a natural and probable consequence, which in the light of the circumstances, should have been foreseen or anticipated." *Bader v. Johnson*, 732 N.E.2d 1212, 1218 (Ind. 2000). . . . The sum of all this is that, in order to be liable for a plaintiff's injury, the harm must have been reasonably foreseeable by the defendant . . . .[89]

If the harm is not reasonably foreseeable, then the defendant is not liable. But as the same court pointed out, an important aspect of foreseeability in Indiana negligence law is "[w]hether the resulting harm is 'foreseeable' such that liability may be imposed on the original wrongdoer is a question of fact for a jury."[90]

Foreseeability is in most cases an issue for the trier of fact, not an issue for the judge to resolve as a matter of law. It becomes an issue for the trial judge only if, as in *Webb* itself, one may conclude that proximate cause is lacking as a matter of law. Under this view the result in *Webb* would remain unchanged because the court determined that, as a matter of law, the injuries to the plaintiff were not reasonably foreseeable. Stated otherwise, the injury to the brother-in-law was not within the scope of liability for over-prescription because the injury was not the harm whose risk made the over-prescription tortious. In sum, layering a foreseeability component onto duty ultimately adds nothing to the analysis and confuses the determination of proximate cause and the reasonableness of an act.

C. The Harm in This Confusion

The interdependence of relationship, foreseeability and duty becomes problematic because it generates the potential for incorrectly understood

---

88. 575 N.E.2d at 997.

89. Control Techniques v. Johnson, 762 N.E.2d 104, 108 (Ind. 2002).

90. *Id.* at 107; *cf.* Shambaugh & Son, Inc. v. Carlisle, 763 N.E.2d 459, 461 (Ind. 2002) (stating "[o]rdinarily, the issue of proximate cause is not properly resolved by summary judgment, but is better left to the jury.").

precedents, which in turn lead to error in subsequent cases. Posit a case where the "real" reason why a defendant prevails on a summary judgment motion is that the defendant acted reasonably under the circumstances, i.e., was not negligent. A modest twist of the facts in *King v. Northeast Security, Inc.*,[91] provides such an example. If a school district hires a respectable security service to oversee its facilities, and arranges for surveillance during times students may be expected to be on its premises, it has acted reasonably to provide a secure environment for its students and should not be liable for a flaw in execution by the service that contributes to a mugging at the school. But if we express that result as a lack of duty to the students, rather than absence of an unreasonable act, it risks suggesting no liability under any circumstances for failure to provide a safe environment.

Expressing the defendant's exposure in terms of duty can overstate liability as well as understate it. Describing a duty in terms of specific circumstances rather than a generalized duty of reasonable care can lead to black-letter rules of liability that approach strict liability. Thus, we may properly say it is unreasonable to fail to do X under these circumstances. But if we formulate that result as "there is a duty to do X," it smacks of strict liability to anyone injured, however remotely.

IV. REFORMULATING THE NEGLIGENCE ACTION

The thesis of this article is that a clear understanding of the issues presented by a negligence case is often frustrated by the need to express a defendant's exposure to liability in terms of "duty." It seems odd to state that the very same action (leaving a banana peel on the marble floor of one's home) is or is not a breach of duty, depending on who gets hurt (the trespasser or the invitee). A more accurate way to describe this legal result would seem to be that it is not reasonable conduct to leave the banana on the floor, but for good reasons we are content to leave the trespasser uncompensated. In the terms recently used in *Mangold*, public policy dictates leaving trespassers where they find themselves as a result of acts of simple negligence on the part of their victims. Similarly, if an injury is beyond the capability of the defendant to prevent at reasonable cost, we should either regard this as immunity (e.g., for governmental failure to prevent crime), or no unreasonable conduct (in a shopkeeper's failure to provide security to patrons from the wholly random holdup). These results amount to a rule of law that this class of plaintiffs cannot recover from this class of defendants under these circumstances. But it is not useful to think of these as the absence of a duty.

A. Duty as a Misleading Term

A major reason to abandon speaking of "duty" is the confusion it generates.

---

91. 790 N.E.2d 474, 477 (Ind. 2003). This case involved a claim against a security agency hired by a school for injuries incurred by a student who was beaten by another student in the school's parking lot.

In ordinary English, most non-lawyers would think the government has a duty to provide law enforcement, and the shopkeeper has a duty to take reasonable steps to provide a safe environment for the shop's patrons. The law should view the matter no differently. If we conclude as a matter of law that there is no liability in these two circumstances, it is not for lack of duty. Rather, that result is based on either a policy that the government should not incur the expense necessary to prevent all crime—the National Guard on every corner—or a conclusion that the shopkeeper's security arrangements were reasonable given the crime rate in the area, available resources, and whatever else might be deemed relevant to the assessment.

To be sure, the duty of the government or the shopkeeper may be expressed as a duty to try to provide security, but not a duty to succeed in that effort. But if that is the test, the concept of duty adds nothing to the concept of liability for unreasonable conduct. It is not a breach of duty to fail to succeed in providing security. It is a breach only if no reasonable steps are taken. Thus, putting it in terms of "duty" and "breach" becomes circular. It adds nothing to the proposition that failure to exercise reasonable care can expose one to liability. More importantly, expressing the legal result in terms of duty is confusing to the fact-finder in many cases. Speaking of no duty to act is understandable if the claim is that the defendant failed to rescue the plaintiff. A failure to take affirmative action does indeed seem wrong if one has a duty to act and perfectly acceptable if there is no duty. But it is not consistent with ordinary language to apply "no duty" to the landlord who abandoned the banana peel. The same is true of many other defendants whom the law protects from liability despite their unreasonable acts.

B. The Role of Policy Considerations and Relationships

If a defendant has acted unreasonably, but the law denies the plaintiff recovery, that result may indeed be based on the relationship between the parties. But the "relationship" factor is no more than a policy consideration that we deem persuasive. In one set of cases, the term "duty" is given a meaning in conformity with its ordinary usage. The common law starts from the proposition that one has no duty to prevent harm to another in the absence of some special relationship between the two. Thus, there is no claim against a bystander who fails to save a drowning person.[92] But a basis for requiring others to take affirmative action may be found in a variety of relationships between the victim and others who are not the immediate cause of the danger.[93] Many of these have already been discussed (landowner—invitee, attorney—client, nursing home–patient, etc.). In those cases, it is quite proper to speak of the relationship's giving rise to a duty, which is another way of saying the defendant may be liable for failing to prevent injury. Finding such a relationship is, I suggest, simply another example of identifying a reason why the law should allow recovery. It represents a policy call that the

---

92. L.S. Ayres & Co. v. Hicks, 40 N.E.2d 334 (1942).

93. *See* J.A.W. v. Roberts, 627 N.E.2d 802 (Ind. Ct. App. 1994).

interests furthered by encouraging affirmative intervention are sufficient to overcome the general rule of nonliability for failure to act. And that in turn depends upon whether we regard the failure to intervene as reasonable conduct on the part of an actor playing the defendant's role under the circumstances.

For example, a legal malpractice claim is, under current doctrine, generally limited to the lawyer's clients.[94] This rule is said to be based on the relationship between the parties. It was presumably grounded in a concern for potentially overwhelming exposure from relatively minor culpability.[95] Yet we already see a shift toward allowing those outside the scope of that relationship to recover for reasonably foreseeable harm inflicted by the attorney.[96] At their core, situations like these are simply matters of policy, i.e., a determination that a certain class of plaintiffs will or will not be permitted to recover for injuries caused by acts of a certain class of defendants that fall below a reasonable standard of conduct. In those cases, the relationship factors balanced in *Webb* boil down to a subset of the public policy considerations already said to be balanced under the *Webb* formulation. *Webb* relied on an Indiana Court of Appeals case for the proposition that "[t]he duty of reasonable care is not, of course, owed to the world at large, but rather to those who might reasonably be foreseen as being subject to injury by the breach of the duty."[97] This is the sentiment expressed by some commentators who contend that the conclusion that everyone is obliged to act reasonably toward everyone, instead of attempting to analyze whether a particular defendant owed a duty to a particular plaintiff, means that "duty in its primary sense necessarily disappears from negligence."[98] This is the result that drove the court of appeals to bifurcate foreseeability. The disappearance of duty is not a result to be avoided. Rather, if properly understood it neither broadens nor narrows exposure to liability because the foreseeability component remains intact in the proof of causation and the reasons for finding duty vel non are equally viable as policy considerations.

Those commentators suggest, and *Webb* stated as much, that "duty" is important to the determination of whether the defendant had an obligation to avoid harming a particular person or class of persons. Although the formulation for a negligence action stated below does not use the word "duty," it still incorporates the same principle, and does so in a clearer line of reasoning.

---

94. Hacker v. Holland, 570 N.E.2d 951, 955 (Ind. Ct. App. 1991) (holding plaintiff must have employed the attorney).

95. Craig D. Martin, *Liability of Attorneys to Non-Clients: When Does a Duty to Non-Clients Exist*, 23 J. LEGAL PROF. 273, 273 (1998).

96. *See, e.g.*, Walker v. Lawson, 526 N.E.2d 968 (Ind. 1988) (allowing action by beneficiary of will who was known third-party beneficiary).

97. *Webb*, 575 N.E.2d at 997 (citing Thiele v. Faygo Beverage, Inc., 489 N.E.2d 562, 574 n.4 (Ind. Ct. App. 1986)).

98. John C.P. Goldberg & Benjamin C. Zipursky, *The Restatement (Third) and the Place of Duty in Negligence Law*, 54 VAND. L. REV. 657, 706 (2001).

C. Limitations on Recovery

Eliminating duty does not expand liability. First, the proximate cause inquiry, undertaken by the jury under current Indiana instructions, asks whether the particular injury to the particular plaintiff was a natural and probable consequence reasonably foreseeable by the defendant. This is Indiana's current formulation of the scope of liability issue. Some would prefer expressing it as whether the plaintiff's injury was caused by the risk that renders the defendant's conduct unreasonable. In either case the focus is not on what the defendant did, but whether the injury should be compensable. Second, many considerations of public policy, including those based on the relationship between the parties, also remain to address whether the defendant, though acting unreasonably, nevertheless avoids liability as a matter of law. Third, if foreseeability or reasonableness of conduct are so clear that they may be addressed as a matter of law, then the trial court may so rule.[99] This analysis reaches the same answers to the "important questions about duty in the primary sense."[100] It simply asks the questions differently, and in my view more lucidly. The effect of jettisoning duty is not a different result in any specific case. It is a better understanding of the principles underlying the result, and therefore more coherent precedent for the future.

D. The Role of Precedent

I suggest the foregoing is a clearer way to think about the question whether a given set of facts supports recovery by the plaintiff against the defendant. But clarity of analysis is not the only value in the law. Predictability and consistency embodied in stare decisis are entitled to a great deal of weight, although some of these cases might be questioned on other substantive grounds. I do not suggest that all of the "rules" expressed in terms of "duty" are no longer good law for failure to formulate the reasoning underlying the rules of liability as I would prefer. Accordingly, although the decisions cited in footnotes six through twenty-two are couched in terms of duty, I do not challenge the result in those cases for that reason. To the extent any of them hold that a given class of defendants does or does not owe a duty to a given class of plaintiffs under the circumstances of that case, that "rule" may equally be reformulated as there is or is not a stated or unstated policy ground why, as a matter of law, those plaintiffs are barred from recovery from those defendants under those circumstances.

For these reasons, then, I would restate the application of the duty concept and, in the process, the broader formulation for a cause of action based on negligence. Essentially, there are three components of a cause of action for negligence, and the defendant may succeed by prevailing on any one of them, in any sequence. First, the plaintiff must prove the defendant was in fact "negligent." Negligence is defined as "[t]he failure to exercise the standard of care that a reasonably prudent person would have exercised in a similar

---

99. *See, e.g.*, Collins v. J.A. House, Inc., 705 N.E.2d 568, 573 (Ind. Ct. App. 1999).

100. Goldberg & Zipursky, *supra* note 98, at 706.

situation."[101] All parties should act with reasonable prudence at all times, and a party who fails to adhere to that standard is negligent. Proof of negligence amounts to a showing of what a reasonably prudent person in the defendant's position would have done, and the failure of the defendant to do so. But merely proving a defendant's negligence does not mean the defendant will necessarily be held liable. A second component of the cause of action requires proof that the defendant's negligence was the proximate cause of actionable harm. In the terms currently used by Indiana courts, that proof consists of a showing that the defendant was both the cause in fact of the harm, and also that the harm was reasonably foreseeable by the defendant at the time of his or her negligence. A plaintiff who provides evidence of negligence, proximate causation and harm has made out the prima facie cause of action. Finally, a defendant is free to contend—on motion for summary judgment, directed verdict or otherwise—that there are policy reasons why the person suffering the harm should nevertheless be precluded from recovery. These reasons may preclude liability to anyone for the particular act or omission, or they may be directed toward barring claims by persons with the characteristics of the particular plaintiff, or recovery for the harm alleged by the plaintiff. In this sense, the original core of the duty concept remains, but it is framed in terms of reasons for allowing or precluding the possibility of recovery.

A majority of the court of which I am a member has not adopted the view of negligence law that I suggest. So long as that remains the case, I see no point to writing separately in judicial opinions as to methodology. In the first place, because *Webb* is existing precedent, the parties usually brief their cases in *Webb* terms, and no one argues for the approach I suggest. Moreover, as already noted, the result in a given case is usually unaffected by choice of methodology. If I agree with the conclusion that the law does or does not permit the plaintiff to recover from the defendant under the circumstances, and the methodology of the opinion is consistent with existing precedent, I expect to concur without elaborating the points made in this Article.[102]

CONCLUSION

In sum, I think the traditional formula of duty, breach, causation and harm is in most cases better understood as proceeding on the assumption that all of us are obliged to take reasonable steps to avoid harm to others in the activities we undertake and can control. The issue of "duty" then resolves itself to an inquiry into whether there is some reason in policy why the law should nevertheless preclude recovery. That reason may arise from, inter alia, the nature of the plaintiff, the nature of the defendant, the relationship between them, the nature of the activities giving rise to the claim, or the nature of the harm alleged.

---

101. BLACK'S LAW DICTIONARY 1056 (7th ed. 1999).

102. *See, e.g.*, Estate of Heck v. Stoffer, 786 N.E.2d 265 (Ind. 2003).

TAKING ETHICAL DISCRETION SERIOUSLY: ETHICAL DELIBERATION AS ETHICAL OBLIGATION

SAMUEL J. LEVINE*

TABLE OF CONTENTS

INTRODUCTION

In recent years, a number of leading scholars have depicted and criticized prevailing models of legal ethics as curtailing the range of ethical discretion and deliberation available to lawyers.[1] These scholars have faulted what they have called, variously but in a nearly uniformly pejorative manner, the "Dominant View,"[2] the "Official View,"[3] the "standard conception of lawyering,"[4] the "regulatory model,"[5] the "traditional professional position,"[6] "the principle of professionalism,"[7] "the libertarian approach,"[8] and the "accepted dogma,"[9]

---

* Associate Professor of Law, Pepperdine University School of Law; LL.M., Columbia University; J.D., Fordham University; Ordination, Yeshiva University; B.A., Yeshiva University. I thank Fraida Liba, Yehudah Tzvi, and Aryeh Shalom for their encouragement.

1. *See, e.g.*, WILLIAM H. SIMON, THE PRACTICE OF JUSTICE: A THEORY OF LAWYERS' ETHICS (1998); Heidi Li Feldman, *Codes and Virtues: Can Good Lawyers Be Good Ethical Deliberators?*, 69 S. CAL. L. REV. 885 (1996); Maura Strassberg, *Taking Ethics Seriously, Beyond Positivist Jurisprudence in Legal Ethics*, 80 IOWA L. REV. 901 (1995); Fred C. Zacharias, *Reconciling Professionalism and Client Interest*, 36 WM. & MARY L. REV. 1303 (1995).

2. SIMON, *supra* note 1, *passim*.

3. David Luban, *The Adversary System Excuse*, *in* THE GOOD LAWYER: LAWYERS' ROLES AND LAWYERS' ETHICS 84 (David Luban ed., 1983).

4. Gerald J. Postema, *Moral Responsibility in Professional Ethics*, 55 N.Y.U. L. REV. 63 (1980).

5. W. Bradley Wendel, *Public Values and Professional Responsibility*, 75 NOTRE DAME L. REV. 1, 8 (1999).

6. Michael Bayles, *Clients and Others*, *in* PROFITS AND PROFESSIONS: ESSAYS IN BUSINESS AND PROFESSIONAL ETHICS 65 (1983), *cited in* Ted Schneyer, *Moral Philosophy's Standard Misconception of Legal Ethics*, 1984 WIS. L. REV. 1529, 1534 n.24.

7. Murray L. Schwartz, *The Professionalism and Accountability of Lawyers*, 66 CAL. L.

REV. 669, 673 (1978). The use of the term "professionalism" in this context is a somewhat striking expression of the wide variation of meanings—at times contradictory—attributed to the term. *Cf.* Monroe H. Freedman, *Professionalism in the American Adversary System*, 41 EMORY L.J. 467, 470 (1992) (arguing that "professionalism means that a lawyer should . . . zealously and competently use all lawful means to protect and advance the client's lawful interests"), *and* Schwartz, *supra*, at 673 (stating that under the principles of professionalism, "[w]hen acting as an advocate, a lawyer must, within the established constraints upon professional behavior, maximize the likelihood the client will prevail"), *with* Anthony T. Kronman, *Legal Professionalism*, 27 FLA. ST. U. L. REV. 1, 5 (1999) (describing legal professionalism as characterized, in part, by a "division of allegiances" between "a particular client" and "the well being of the law as a whole"), *and* Zacharias, *supra* note 1, at 1307 (stating that professionalism "encompasses the notion that the lawyer's function includes . . . the ability to distance oneself from personal and client desires in order to evaluate the effect of potential actions on clients, third parties, and the legal system"). *See generally* Samuel J. Levine, *Faith in Legal Professionalism: Believers and Heretics*, 61 MD. L. REV. 217 (2002).

For an example of these variations in the context of the Jewish identity of some lawyers and, more broadly, group identity, contrast Sanford Levinson, *Identifying the Jewish Lawyer: Reflections on the Construction of Professional Identity*, 14 CARDOZO L. REV. 1577, 1578-79 (1993), describing "one important aspect of ([one] version of) the professional project" as "the 'bleaching out' of merely contingent aspects of the self, including the residue of particularistic socialization that we refer to as our 'conscience'" and concluding that "[t]he triumph" of the "standard version of the professional project" would result in "almost purely fungible members of the respective professional community" through which "[s]uch apparent aspects of the self as one's race, gender, religion, or ethnic background would become irrelevant to defining one's capacities as a lawyer" with Russell G. Pearce, *Jewish Lawyering in a Multicultural Society: A Midrash on Levinson*, 14 CARDOZO L. REV. 1613, 1635, 1636 (1993) (quoting MARTHA MINOW, MAKING ALL THE DIFFERENCE: INCLUSION, EXCLUSION, AND AMERICAN LAW 376 (1990)), stating that the legal profession must acknowledge the "contradiction between group identity and the goals of the professional project" and should thus "discard[] the notion of 'bleaching out'" in favor of "seek[ing] to 'create community' by bringing us together to explore the potential for rule of law in light of 'how we are all different from one another and also how we are all the same'". *See generally* Samuel J. Levine, *The Broad Life of the Jewish Lawyer: Integrating Spirituality, Scholarship, and Profession*, 27 TEX. TECH L. REV. 1199 (1996) [hereinafter Levine, *Broad Life*]; Samuel J. Levine, *Professionalism Without Parochialism: Julius Cohen, Rabbi Nachman of Breslov, and the Stories of Two Sons*, 71 FORDHAM L. REV. 1339 (2003) [hereinafter Levine, *Professionalism*]; Russell G. Pearce, *The Jewish Lawyer's Question*, 27 TEX. TECH L. REV. 1259 (1996).

Indeed, the inevitable tension that arises when different conceptions of professionalism are considered in light of religious and personal values has been a central focus of a "religious lawyering movement." *See* Russell G. Pearce, *Foreword: The Religious Lawyering Movement: An Emerging Force in Legal Ethics and Professionalism*, 66 FORDHAM L. REV. 1075 (1998); *see also* Samuel J. Levine, *Introductory Note: Responding to the Problems of Ethical Schizophrenia*, 38 CATH. LAW. 145 (1998); Symposium: *Faith and the Law*, 27 TEX. TECH L. REV. 911 (1996); Symposium: *Lawyering and Personal Values*, 38 CATH. LAW. 145 (1998); Symposium, *Rediscovering the Role of Religion in the Lives of Lawyers and Those They Represent*, 26 FORDHAM URB. L.J. 821 (1999); Symposium, *The Relevance of Religion to a Lawyer's Work: An*

characterized by adherence to "technocratic lawyering,"[10] "categorical" ethical decisionmaking,[11] "positivism,"[12] "conventionalism,"[13] "role-differentiation,"[14] and the "ideology of advocacy."[15] Although these criticisms are far from identical, they share the general premise that the ethical nature of legal practice suffers from literalistic adherence to what appears to be the letter of ethics codes, combined with overzealous loyalty to clients' wishes, obviating the need for—or, at times, preventing the possibility of—careful attention to ethical issues.[16]

In response, many of these scholars have proposed alternative models of legal ethics, with titles such as the "Contextual View,"[17] the "Discretionary Model,"[18] "Interpretive Integrity,"[19] the "Principle of Integrative Positivism,"[20] the "Integrity Thesis"[21] and a "common law of lawyers' ethics,"[22] each aimed at

---

Interfaith Conference, 66 FORDHAM L. REV. 1075 (1998).

8. William H. Simon, *Ethical Discretion in Lawyering*, 101 HARV. L. REV. 1083, 1085 (1988).

9. ALAN H. GOLDMAN, THE MORAL FOUNDATIONS OF PROFESSIONAL ETHICS 90 (1980), *cited in* Schneyer, *supra* note 6, at 1534 n.24.

10. Feldman, *supra* note 1, *passim*; Wendel, *supra* note 5, at 8.

11. SIMON, *supra* note 1, at 9.

12. Strassberg, *supra* note 1, *passim*.

13. Wendel, *supra* note 5, at 17.

14. Postema, *supra* note 4, at 63.

15. William H. Simon, *The Ideology of Advocacy, Procedural Justice and Professional Ethics*, 1978 WIS. L. REV. 29.

16. Professor Rob Atkinson has described as "reformers" those scholars, including himself, who "believe that other normative limits, sometimes narrower than the letter of the law, govern what the good person, as good lawyer, may do for at least some clients, at least some of the time." Rob Atkinson, *A Skeptical Answer to Edmunson's Contextualism: What We Know We Lawyers Know*, 30 FLA. ST. U. L. REV. 25, 26-27 (2002). *Cf.* David Rosenthal, *The Criminal Defense Attorney, Ethics, and Maintaining Client Confidentiality: A Proposal to Amend Rule 1.6 of the Model Rules of Professional Conduct*, 6 ST. THOMAS L. REV. 153, 157 (1993) (applying the title "Moralists" to theorists who aim to "promote the lawyer's responsibility to seek societal justice over the obligation to assist a particular client").

In one of the earliest and most powerful responses to such a position, Professor Ted Schneyer employed the largely accurate term "philosophers" to describe many of these scholars and attributed to them a "standard misconception of legal ethics." Schneyer, *supra* note 6; *see also* Ted Schneyer, *Some Sympathy for the Hired Gun*, 41 J. LEGAL EDUC. 11 (1991); Zacharias, *supra* note 1, at 1326-27 (describing a "relatively new branch of philosophical scholarship [that] concentrates so much on perceived client orientation of lawyers that it suggests that objectivity in legal practice is virtually non-existent" and concluding that "this school of thought may overstate the case").

17. SIMON, *supra* note 1, at 9.

18. *Id.* at 10.

19. Strassberg, *supra* note 1, at 934.

20. Serena Stier, *Legal Ethics: The Integrity Thesis*, 52 OHIO ST. L.J. 551 (1991).

21. *Id.*

22. Feldman, *supra* note 1, at 945.

providing a framework through which lawyers are encouraged—if not required—to replace simplistic ethical decisionmaking with more thoughtful and complex deliberation of ethical considerations. These models, like the criticisms of the prevailing model, differ substantially from one another, including proposals for rethinking the function and goals of the lawyer,[23] providing civil remedies for ethical violations,[24] suggestions of additional ethics provisions,[25] and innovative methodologies for the interpretation of current ethics codes.[26]

Despite these differences, however, certain common themes run through all of the models. On a practical level, each of the models requires a fundamental change in the way the current regulatory system of ethics codes is interpreted and/or implemented. More importantly, for the purposes of the current analysis, it seems that each of the models either prescribes to the lawyer a particular decision to be applied to a given ethical dilemma or does not obligate the lawyer to demonstrate the exercise of thoughtful ethical discretion. Thus, the models all appear to fall short of their stated goal: ultimately permitting but not requiring that the individual lawyer engage in ethical deliberation and analysis.

This Article looks to build on and respond to the work of these scholars, adopting and explicating many of their insights into the nature of the current structure of legal ethics and ethics codes. After describing[27] and critiquing two representative and prevailing alternative models,[28] the Article proposes a "Deliberative Model" which not only allows for, but expects and, at times, requires careful ethical deliberation prior to the exercise of discretion.[29]

Specifically, the Article posits that the concept of ethical discretion in legal practice should be understood neither as a license to engage in a variety of optional activities without justification for a particular course of action, nor as grounds for mandating a particular decision under circumstances of ethical complexity. Instead, the Article suggests that the lawyer's professional responsibility carries with it a duty on the individual lawyer to exercise such discretion through consideration of the relevant ethical issues. Thus, the Article takes seriously the principle of ethical discretion, respecting the role of individual ethical decisionmaking, but requiring that such decisionmaking be carried out through a justifiable process of ethical deliberation.

---

23. SIMON, *supra* note 1, *passim*.

24. Feldman, *supra* note 1, at 945-47.

25. *See, e.g.*, Russell G. Pearce, *Model Rule 1.0: Lawyers are Morally Accountable*, 70 FORDHAM L. REV. 1805, 1805 (2002); Zacharias, *supra* note 1.

26. Stier, *supra* note 20; Strassberg, *supra* note 1.

27. *See infra* Part I.

28. *See infra* Part II.

29. *See infra* Part III.

I. REPRESENTATIVE AND PREVAILING ALTERNATIVE MODELS OF LEGAL ETHICS

A. Simon's "Contextual Model"

Among the work of scholars criticizing the current state of ethics regulation, Professor William Simon's may stand out as both the most prominent and the most ambitious. In 1998, Simon published *The Practice of Justice*, a book-length elaboration of some of the theories he had begun to develop in a series of law review articles.[30] The target of Simon's critique is the "Dominant View," his appellation, for "[t]he prevailing approach to lawyers' ethics as reflected in the bar's disciplinary codes, the case law on lawyer discipline, and the burgeoning commentary on professional responsibility."[31] Under the Dominant View, "the lawyer must—or at least may—pursue any goal of the client through any arguably legal course of action and assert any nonfrivolous legal claim."[32] According to Simon, "in the Dominant View the only ethical duty distinctive to the lawyer's role is loyalty to the client. Legal ethics impose no responsibilities to third parties or the public different from that of the minimal compliance with law that is required of everyone."[33] Simon sees this emphasis on loyalty to the client as central to both the Model Code of Professional Responsibility and the Model Rules of Professional Conduct, reflected in provisions promoting zealous advocacy[34] and confidentiality.[35]

Simon finds the Dominant View and the accompanying regulatory scheme insufficient as a basis for legal ethics, criticizing both the form of current ethics codes and, consequentially, their mode of interpretation. Simon sees the form of the codes—in particular the Model Rules—as "a set of mechanical disciplinary rules" that, he concludes, "reduce ethics to a matter of mindless rule application."[36] Specifically, he argues, the Dominant View imposes a form of

---

30. *See* SIMON, *supra* note 1, at 249 (acknowledging the book's reliance, in parts, on: Simon, *supra* note 8; William H. Simon, *Should Lawyers Obey the Law?*, 38 WM. & MARY L. REV. 217 (1996); William H. Simon, *The Ethics of Criminal Defense*, 91 MICH. L. REV. 1703 (1993)).

31. SIMON, *supra* note 1, at 7.

32. *Id.* at 7.

33. *Id.* at 8.

34. *See* MODEL CODE OF PROFESSIONAL RESPONSIBILITY [hereinafter MODEL CODE], Canon 7 ("A lawyer shall represent a client zealously within the bounds of the law."); ABA MODEL RULES OF PROFESSIONAL CONDUCT [hereinafter MODEL RULES], R. 1.3, cmt. [1] ("A lawyer must also act with commitment and dedication to the interests of a client and with zeal in advocacy upon the client's behalf.").

35. *See* MODEL CODE, *supra* note 34, DR 4-101; MODEL RULES, *supra* note 34, R. 1.6.

36. SIMON, *supra* note 1, at 15. As Simon observes, the Model Rules generally employ the form of "black letter rules" in place of the Model Codes' inclusion of aspirational "Ethical Obligations" alongside more specific and mandatory "Disciplinary Rules." *Id.* at 14-15. Indeed, scholars have documented the evolution of twentieth century ethics regulations from "fraternal norms issuing from an autonomous professional society" to "binding legal rules . . . rendered in

"categorical" ethical decisionmaking, "severely restrict[ing] the range of considerations the decisionmaker may take into account when she confronts a particular problem."[37] In short, "a rigid rule dictates a particular response in the presence of a small number of factors."[38] As a result, "[t]he decisionmaker has no discretion to consider factors that are not specified or to evaluate specified [sic] factors in ways other than those prescribed by the rule[s]."[39] Ultimately, Simon rejects the current model of ethics regulation, stating poignantly that "it is no longer apparent what it has to do with ethics or responsibility."[40]

Simon is particularly troubled by the jurisprudential implications of categorical—or "formalistic"—modes of ethical interpretation.[41] He notes, with more displeasure than irony,[42] that "[t]he revolt against formalism in legal

---

statutory language." Geoffrey C. Hazard, Jr., *The Future of Legal Ethics*, 100 YALE L.J. 1239, 1254, 1250-51 (1991). *See also* CHARLES W. WOLFRAM, MODERN LEGAL ETHICS 53-63 (1986); Geoffrey C. Hazard, Jr., *Legal Ethics: Legal Rules and Professional Aspirations*, 30 CLEV. ST. L. REV. 571 (1982); Geoffrey C. Hazard, Jr., *The Legal and Ethical Position of the Code of Professional Ethics*, *in* 5 SOCIAL RESPONSIBILITY: JOURNALISM, LAW, MEDICINE 5 (Louis Hodges ed., 1979) [hereinafter Hazard, *Legal and Ethical*]; Samuel J. Levine, *Taking Ethics Codes Seriously: Broad Ethics Provisions and Unenumerated Ethical Obligations in a Comparative Hermeneutic Framework*, 77 TULANE L. REV. 527 (2003).

These changes have been met with sharply differing reactions among scholars, reflecting starkly opposing attitudes toward both the form and function of ethics rules. Like many of the scholars whose work is highlighted in this Article, Simon faults what he sees as the overly "simple and straightforward" nature of ethics codes. SIMON, *supra* note 1, at 14. Conversely, other scholars, who approve of the increasing specificity of the codes, have expressed their own dissatisfaction with areas of the codes that, in their view, have remained unacceptably broad and imprecise. *See* Levine, *supra*. In fact, as a result of these criticisms, the quest for "black-letter law" regulating the conduct of lawyers has continued, culminating in the completion in the fall of 2000 of the American Law Institute's RESTATEMENT OF THE LAW GOVERNING LAWYERS [hereinafter RESTATEMENT].

37. SIMON, *supra* note 1, at 9.

38. *Id.*

39. *Id.*

40. *Id.* at 15.

41. *Id.* at 3.

42. At times, irony seems an appropriate reaction to anomalous and hypocritical conduct on the part of individual lawyers and the organized bar. *See* Thomas L. Shaffer, *The Irony of Lawyers' Justice In America*, 70 FORDHAM L. REV. 1857 (2002).

Simon is similarly unforgiving in his attack on the bar's approach to confidentiality, consciously engaging in an "*ad hominem* argument[]" because "the indications of bad faith" on the part of the bar "are too salient to pass over." SIMON, *supra* note 1, at 56. Simon characterizes the bar's defenses of confidentiality as "sloppy, cavalier, and dogmatic," *id.*; he describes the rationale offered in the Model Rules for confidentiality as "ludicrously inconsistent" with the substance of the rule, *id.*; and he offers illustrations of "the bar's anxiety and hypocrisy about its confidentiality norms." *Id.* at 222-23 n.9. *Cf.* MONROE H. FREEDMAN & ABBE SMITH, UNDERSTANDING LAWYERS' ETHICS 144 (2d ed. 2002) (stating that "[s]ome of the ABA's exceptions to lawyer-client

thought has carried the day in nearly every legal field other than that of lawyering itself."[43] Simon finds it "remarkable" that "[a]lone . . . professional responsibility . . . ha[s] enjoyed relative immunity from the[ ] critiques" of formalism that have been applied by modern jurisprudence "nearly everywhere else."[44]

Simon offers an alternative model of ethical decisionmaking that embraces complexity. He calls it the "Contextual View."[45] As the name suggests, the "essence of this approach is contextual judgment," which Simon defines as "a judgment that applies relatively abstract norms to a broad range of the particulars [that apply in the] case at hand."[46] The "basic maxim" of this approach is indeed

---

confidentiality are a mockery of an ideal . . . .").

43. SIMON, *supra* note 1, at 3. Professor Deborah Rhode has similarly faulted lawyers for acting in a "highly selective" manner when objecting to the application of "vague" terms such as "justice," "fairness," and "good faith" for imposition of moral condemnation or discipline. DEBORAH L. RHODE, IN THE INTERESTS OF JUSTICE: REFORMING THE LEGAL PROFESSION 68 (2000). As Professor Rhode notes, "[t]he legal system routinely requires judges, juries, and prosecutors to pursue 'justice' or to determine 'fairness,' and it imposes sanctions on businesses that fail to act in 'good faith.'" *Id.* Echoing Simon in emphasizing the "selectiv[ity]"—if not downright hypocrisy—of arguments against applying these concepts to the conduct of lawyers as well, Rhode concludes with the observation that "[l]awyers charge substantial fees for interpreting such requirements. The interpretive process is no different when lawyers' own actions are involved." *Id.*

44. SIMON, *supra* note 1, at 3. Likewise, Simon connects the Dominant View with the jurisprudential philosophy of Positivism to the extent that it "characteriz[es] the values that compete with client loyalty . . . as nonlegal." *Id.* at 17. Here, too, however, he observes that "strong versions of Positivism turn out to be implausible, and indeed they are rejected by most lawyers outside the sphere of legal ethics." *Id.* at 17-18. *See also id.* at 37 ("While a few legal philosophers still defend the Positivist premise, nearly all practicing lawyers reject it implicitly in the way they argue cases, advise clients, and draft documents. Legal ethics is the only area in which they continue to cling to it."). As Professor Bradley W. Wendel has put it, "[t]he regulatory model is a jurisprudential dinosaur." Wendel, *supra* note 5, at 17. *See also id.* (stating that "the regulatory model assumes that the legal rules may be applied mechanically, without resort to creative normative judgment[, and] [i]n this way, it resembles a jurisprudential position that has almost no adherents outside the realm of legal ethics").

Clearly, though largely implicitly, Simon's analysis relies in part on insights developed by the Legal Realist and Critical Legal Studies movements. Although the substance of his arguments contains few citations to specific works or ideas in these fields, Simon acknowledges Critical Legal Studies as one of two inspirations for his argument. SIMON, *supra* note 1, at 247. In addition, a relatively brief list of recommended "Further Reading" at the end of the book includes citations to such works. *See id.* at 243-46 (citing, e.g., Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685 (1976); David B. Wilkins, *Legal Realism for Lawyers*, 104 HARV. L. REV. 468 (1990)).

45. SIMON, *supra* note 1, at 9. Simon also refers to this as the "Discretionary Model." *Id.* at 10.

46. *Id.* at 10.

complex and requires that the lawyer "take such actions as, considering the relevant circumstances of the particular case, seem likely to promote justice."[47] Under such an approach, Simon asserts, ethical decisions will "often turn on 'the underlying merits'" of an issue.[48]

Simon readily acknowledges—in fact, he relies in part upon—the similarity between the model he proposes for lawyers generally and that which is most familiarly associated with the prosecutor. Simon writes that his formulation of the basic maxim of the Contextual View was "partly inspired" by the Model Code's prescription for the role of the prosecutor "to seek justice, not merely to convict."[49] At first glance, it seems striking that Simon adopts so vague a command as "seek justice" to guide lawyers' ethical decisionmaking. After all, while there exists a wide-ranging debate regarding the appropriate level of specificity for ethics guidelines,[50] even scholars such as Professor Fred Zacharias, who accept some measure of ambiguity in ethics provisions,[51] find the

---

47. *Id.* at 9. In adopting a position that obligates a private lawyer to do justice, Simon participates in a debate that has continued "at least since the rise of large-scale organizations in American society during the nineteenth century." Susan D. Carle, *Lawyers' Duty to Do Justice: A New Look at the History of the 1908 Canons*, 24 LAW & SOC. INQUIRY 1, 2 (1999). *See id.* ("Perhaps no issue in legal ethics has been debated more often, and resolved less satisfactorily, than that of lawyers' duties—in civil cases—to concern themselves with the 'justice' of their client's cause.").

48. SIMON, *supra* note 1, at 9.

49. *Id.* at 10. *See* MODEL CODE, *supra* note 34, EC 7-13 (providing that "[t]he responsibility of the public prosecutor differs from that of the usual advocate; his duty is to seek justice, not merely to convict."). *See also* MODEL RULES, *supra* note 34, R. 3.8, cmt. [1] (stating the proposition that "[a] prosecutor has the responsibility of a minister of justice and not simply that of an advocate."). For historical outlines of the prosecutor's duty to seek justice, see Steven K. Berenson, *Public Lawyers, Private Values: Can, Should, and Will Government Lawyers Serve the Public Interest?*, 41 B.C. L. REV. 789, 792-94 (2000) and Bruce A. Green, *Why Should Prosecutors "Seek Justice"?*, 26 FORDHAM URB. L.J. 607, 612-17 (1999).

50. *See supra* note 36.

51. *See* Fred C. Zacharias, *Foreword: The Quest for a Perfect Code*, 11 GEO. J. LEGAL ETHICS 787, 791 (1998) (stating that "I am not arguing that codes need to be specific in every regard"); *id.* at 791-92 n.44 (stating that "I have taken a contrary position, in sharp contrast to the position of scholars . . . who favor specificity in regulation whenever it is possible"); *id.* (contrasting Fred C. Zacharias, *Specificity in Professional Responsibility Codes: Theory, Practice, and the Paradigm of Prosecutorial Ethics*, 69 NOTRE DAME L. REV. 223, 249-85 (1993) [hereinafter Zacharias, *Specificity*], with Bruce A. Green & Bernadine Dohrn, *Foreword: Children and the Ethical Practice of Law*, 64 FORDHAM L. REV. 1281, 1282-83, 1288-89 (1996); Bruce A. Green, *Whose Rules of Professional Conduct Should Govern Lawyers in Federal Court and How Should the Rules be Created?*, 64 GEO. WASH. L. REV. 460, 468 (1996); Bruce A. Green, *Zealous Representation Bound: The Intersection of the Ethical Codes and the Criminal Law*, 69 N.C. L. REV. 687, 689-90 (1991)); *see also* Zacharias, *supra* note 1, at 1328 n.85 (stating that "[i]t might be possible to adopt a highly specific code of professional conduct, which all lawyers must obey . . . [that] might instruct lawyers in how to balance competing interests in all situations," but noting

"do justice" standard to be unworkably broad as applied to prosecutors, let alone private attorneys.[52] Nevertheless, Simon's theory of "justice" does not appear vulnerable to such concerns, as he dedicates an entire chapter to a detailed discussion of the considerations underlying the Contextual View, applied to concrete situations.[53]

---

that "[v]irtually all existing codes . . . avoid that approach, probably for good reason").

52. *See* Fred C. Zacharias, *Structuring the Ethics of Prosecutorial Trial Practice: Can Prosecutors Do Justice?*, 44 VAND. L. REV. 45 (1991); Zacharias, *Specificity*, *supra* note 51, at 292 (asserting that: "[f]or obvious reasons . . . the 'justice' approach seems inadequate to reformers, both as a general method of role definition and as a means to alleviate the reformers' specific concerns regarding [prosecutorial misconduct]"; that "[e]mpirically, the justice provisions have not worked"; and that "in the absence of other constraints, a ['do justice'] rule provides minimal behavioral guidance and, in practice, may even reduce prosecutors' ethical introspection"); *cf.* Geoffrey C. Hazard, *Law and Justice in the Twenty-First Century*, 70 FORDHAM L. REV. 1739, 1741 (2002) (asserting that "[i]n contemporary American academic discussion there is much talk about 'justice' that is not anchored in the mundane apparatus of judges and court clerks, pleadings and procedural motions, and the technicalities of legal interpretation," and concluding that "[i]n my view these discussions are vacuous").

But see Berenson, *supra* note 49, at 817 (stating that "the 'do justice' standard [can] serve as a basis for determining the appropriate professional role for public prosecutors despite criticisms regarding its vagueness [and] . . . can also serve as an important source for determining the appropriate professional role for government lawyers in civil litigation contexts"); Green, *supra* note 49, at 634 (finding that "[d]oing justice comprises various objectives which are, for the most part, implicit in our constitutional and statutory schemes"). *Cf.* Catherine J. Lanctot, *The Duty of Zealous Advocacy and the Ethics of the Federal Government Lawyer: The Three Hardest Questions*, 64 S. CAL. L. REV. 951, 957-58 (1991) (noting the "ethical tension inherent in the role of the government lawyer," but concluding that "the ethical codes draw no distinctions between the duty of the government lawyer and the duty of the private lawyer to defend a civil case zealously").

53. *See* SIMON, *supra* note 1, at 138-69; *see also* Ted Schneyer, *Reforming Law Practice in the Pursuit of Justice: The Perils of Privileging "Public" Over Professional Values*, 70 FORDHAM L. REV. 1831, 1842 & n.68 (2002) (reviewing RHODE, *supra* note 43, and stating that "[t]he main problem with Rhode's call for a more 'contextual' view of legal ethics is the difficulty of clarifying how lawyers (or their regulators) should determine the import of contextual variables or how finely-grained their contextual analysis should be," and contrasting Simon's work as a "more sustained effort to identify contextual factors that should be considered relevant").

Nevertheless, a number of scholars have argued that Simon's model fails to provide sufficiently specific guidance. *See, e.g.*, Heidi Li Feldman, *Matter of Ethics Apparently Substantial, Oddly Hollow: The Enigmatic Practice of Justice*, 97 MICH. L. REV. 1472, 1480-81 (1999) (stating that Simon "recommends procedural guidelines to assist lawyers in deciding what justice requires of them in any given situation," but concluding that the guidelines "do not provide a lawyer with a substantive account of justice"); David Luban, *Reason and Passion in Legal Ethics*, 51 STAN. L. REV. 873, 896-97 (1999) (citing SIMON, *supra* note 1, at 157) (quoting Simon's view that "when the lawyer decides that time and resources permit only an incomplete analysis, the lawyer will fall back on 'presumptive responses to broad categories of situations,'—in a word, she will fall back on rules," but concluding that Simon is "silen[t] on the issue of default rules" and that

Moreover, responding to the potential argument that the Contextual View "collapses the lawyer's role into that of the . . . prosecutor,"[54] Simon explains that the comparison is only to the prosecutor's "style of judgment, not the particular decisions . . . prosecutors make."[55] Simon insists that the Contextual View thus "incorporates much of the traditional lawyer role, including the notion that lawyers can serve justice through zealous pursuit of clients' goals" and "is fully respectful of the most plausible conceptions of procedural justice and the adversary system."[56]

One of the primary targets of Simon's critique is the Dominant View's "close-to-absolute confidentiality norms for client information."[57] Of course, as Simon acknowledges, he is far from the first to criticize the rules of attorney-client confidentiality.[58] Yet, in both tone and substance, Simon is particularly relentless in his attack on the organized bar's attitude toward confidentiality,[59] as reflected in ethics codes.[60] In fact, the very first example Simon offers to

---

the Contextual View "offers no purchase in figuring out what the default rules should be"); Deborah L. Rhode, *Symposium Introduction: In Pursuit of Justice*, 51 STAN. L. REV. 867, 872 (1999) (stating that although Simon "proposes some promising reforms [he] gives little attention to how they might be achieved or to the social, economic, and political barriers that stand in the way"). Indeed, Simon himself refers to "[t]he scant efforts I've made to suggest how my ideas might be institutionalized and implemented." SIMON, *supra* note 1, at 248.

54. SIMON, *supra* note 1, at 10-11.

55. *Id.* at 11. Simon similarly justifies a further comparison between the Contextual View of a lawyer's ethical judgment and the style of contextual judgment associated with judges. *See id.* at 10-11.

56. *Id.* at 11.

57. *Id.* at 54.

58. *See, e.g.*, Daniel R. Fischel, *Lawyers and Confidentiality*, 65 U. CHI. L. REV. 1 (1998); Steven Shavell, *Legal Advice About Contemplated Acts: The Decision to Obtain Advice, Its Social Desirability, and Protection of Confidentiality*, 17 J. LEGAL STUDIES 123 (1988); Fred C. Zacharias, *Rethinking Confidentiality*, 74 IOWA L. REV. 351 (1989).

59. *See supra* note 42.

60. SIMON, *supra* note 1, at 54-62. Professor Susan Koniak has offered an insightful examination of the "centrality and power of the norm of confidentiality in the bar's nomos." Susan P. Koniak, *The Law Between the Bar and the State*, 70 N.C. L. REV. 1389, 1427-47 (1992). Koniak refers to confidentiality as a "constitutional norm" in the bar's nomos, a norm

> so central to group definition . . . that the group perceives threats to the norm as threats against . . . the group's very existence; that the group sees proposals to change the norm as proposals to change the . . . group itself; and consequently that the group feels extreme action in defense of the norm is justified.

Id. at 1427.

Among other observations, Koniak notes that:

> the special importance of the norm of confidentiality in the bar's nomos is not apparent on the face of the ethics codes . . . but the ethics opinions interpreting the codes make it plain. A pattern emerges in these opinions: rules affirming a duty or the discretion not disclose are either narrowed to the point of near-irrelevance or held to be overridden by

contrast the Dominant View with the Contextual View, and one to which he returns at several points in the book, is the vexing case of the "Innocent Convict."[61] Simon uses this term to refer to a situation in which, during a conversation between a lawyer and a client, the client admits to having committed a capital crime for which an innocent individual has been convicted.[62] Under traditional rules of confidentiality, such a lawyer would be prohibited from disclosing this information, even if the Innocent Convict faced execution.[63]

Although, as he again acknowledges, Simon is expressing a widely held view among scholars that the scenario of the Innocent Convict represented an unsatisfactory outcome of the confidentiality rules,[64] Simon's solution serves as

---

rules requiring silence.

Id. at 1428, 1431.

61. SIMON, *supra* note 1, at 4.

62. This scenario has received extensive attention among ethics scholars. Indeed, Professor Wolfram has referred to this scenario as a "much mooted situation." WOLFRAM, *supra* note 36, at 673. *See, e.g.*, Symposium, *Executing the Wrong Person: The Professionals' Ethical Dilemmas*, 29 LOY. L.A. L. REV. 1543 (1996). One leading ethics scholar has stated that "[o]f all the 'doomsday scenarios' commonly discussed in law school classrooms, this one holds special fascination, because it is so stark and so real." W. William Hodes, *What Ought to Be Done—What Can be Done—When the Wrong Person Is in Jail or About to Be Executed? An Invitation to a Multi-Disciplined Inquiry, and a Detour About Law School Pedagogy*, 29 LOY. L.A. L. REV. 1547, 1561 (1996).

63. The obligation to maintain the client's confidence in this scenario is implicit but unambiguous in the Model Code, which permits disclosure of a client's intention to commit a crime, but contains no provision permitting disclosure of past crimes, regardless of the consequences to a third party. *See* MODEL CODE, *supra* note 34, DR 4-101 (C)(3).

Although various drafts of the Model Rules would have permitted, or even mandated, disclosure to prevent various negative consequences to a third party, *see* STEPHEN GILLERS & ROY D. SIMON, REGULATION OF LAWYERS: STATUTES AND STANDARDS 80 (2003), as enacted in 1983, the Model Rules permitted disclosure to prevent a client from committing crimes resulting in death or serious physical injury, but did not provide for disclosure of past crimes. *See id.* at 70.

It should be noted that the current version of the Model Rules, recently amended, permit disclosure "to prevent reasonably certain death or substantial bodily harm," MODEL RULES, *supra* note 34, R. 1.6 (b) (2), apparently permitting—though still not requiring—disclosure to save the Innocent Convict from execution, and presumably from imprisonment as well. *See id.*, cmt. 6 (describing Rule's recognition of "overriding value of life and physical integrity"). *See* discussion *infra* Part II.C; *see also* RESTATEMENT, *supra* note 36, § 66.

64. *See, e.g.*, Zacharias, *supra* note 58. As Simon notes, even "[m]any defenders of the Dominant View regard the case as sufficiently troubling to warrant an exception," leading some to engage in "tortuous analyses of [the] problem" in an "effort to justify an exception without trenching further on the categorical confidentiality norm." SIMON, *supra* note 1, at 218 n.6. *See id.* (citing Symposium, *supra* note 62).

As Simon further observes, the approach to the Innocent Convict in the Model Rules was particularly disturbing when juxtaposed with the conversely categorical exception to confidentiality

to establish a claim or defense on behalf of the lawyer in a controversy between the

a useful example for understanding more broadly his approach to ethical decisionmaking for lawyers. Simon first posits that "the general procedural system has failed" the Innocent Convict and that the system is not likely to "correct itself."[65] Therefore, in accordance with the guidelines he has delineated for the Contextual View, Simon asserts that "[t]he lawyer thus has some responsibility to assess substantive merit."[66] Undoubtedly, as Simon concludes, such an assessment will lead the lawyer to conclude that the execution of the Innocent Convict will produce "a horrendous injustice."[67] The ethical question for the lawyer, then, is how to respond to the prospect of such injustice, in the absence of a tenable interpretation of the Model Rules permitting disclosure.[68]

For Simon, the unequivocal answer is for the lawyer to disclose the confidential information to save the life of the Innocent Convict. To the extent that this answer requires violation of the Model Rules, Simon declares that "the lawyer might have to consider disclosure as a form of nullification."[69] Using harsh language characteristic of his attitude toward rules prescribing nearly categorical adherence to confidentiality,[70] Simon decries the "substantive absurdity of the rules' application in the particular circumstances of the Innocent

---

> lawyer and the client, to establish a defense to a criminal charge or civil claim against the lawyer based upon conduct in which the client was involved, or to respond to allegations in any proceeding concerning the lawyer's representation of the client.

MODEL RULES, *supra* note 34, R. 1.6 (b) (3). *See* SIMON, *supra* note 1, at 222-23 n.9; *see also* Roger C. Cramton & Lori P. Knowles, *Professional Secrecy and its Exceptions:* Spaulding v. Zimmerman *Revisited*, 83 MINN. L. REV. 63 (1998):

> [A] profession that justifiably asks for and receives permission to disclose confidential client information when its own economic interests are at stake (*e.g.* to collect a fee from a client) cannot plausibly take the position that the threatened death or serious injury of another does not justify an occasional sacrifice of confidentiality.

Id. at 111-12; Fischel, *supra* note 58:

> The same lawyer who is prohibited from disclosing information learned while representing a client to exonerate someone falsely accused of a capital crime . . . is perfectly free to disclose confidential information when he or she is the one accused, falsely or not. Nor is there any requirement that the lawyer's liberty be at stake, or even that the lawyer be accused of anything criminal. A simple fee dispute with a client is sufficient grounds to disclose confidential information. The lawyer's interest in collecting a fee is apparently a higher priority than exonerating an innocent defendant about to be convicted of a capital crime . . . . Confidentiality means everything in legal ethics unless lawyers lose money, in which case it means nothing.

Id. at 10. *See supra* note 42.

65. SIMON, *supra* note 1, at 163.
66. *Id.*
67. *Id.*
68. *See supra* note 64.
69. SIMON, *supra* note 1, at 164.
70. *See supra* note 42.

Convict."[71] Such a rule, he continues, would be "degrading" to the lawyer, for whom "it would be grotesque not to disclose" confidences to save the life of the Innocent Convict.[72] Though he offers admittedly perfunctory efforts to reconcile his position with the rules—suggesting, for example, that perhaps "[t]he rule should not be interpreted to require such degradation"[73]—Simon recognizes that such reconciliation is impossible. Therefore, he confronts the rule directly, insisting that "[i]f such an interpretation is unavoidable, then the lawyer should defy the rule."[74]

B. Strassberg's Model of "Interpretive Integrity"

As a result of the innovative and provocative nature of his theories and observations, Simon's work has been the subject of both considerable admiration[75] and substantial criticism.[76] Indeed, among Simon's critics are many scholars who are largely sympathetic to his arguments but who nevertheless maintain reservations about accepting the full range of his proposals.[77]

---

71. SIMON, *supra* note 1, at 164.

72. *Id.*

73. *Id.*

74. *Id.*

75. *See, e.g.*, Review Essay Symposium, *The Practice of Justice by William H. Simon*, 51 STAN. L. REV. 867 (1999); Rob Atkinson, *Lawyering in Law's Republic*, 85 VA. L. REV. 1505, 1517 (1999); Robert F. Cochran, Jr., *The Rule of Law(yers)*, 65 MO. L. REV. 571 (2000); Feldman, *supra* note 53.

76. *See, e.g.*, *supra* note 53 and *infra* note 119.

77. As Professor Deborah Rhode noted in the introduction to a *Stanford Law Review* Symposium dedicated to reviews of Simon's book:

> Like most leading ethics experts, the participants here are all FOBs—friends of Bill—in important respects. They share his premise that the bar's prevailing ethical norms are fundamentally flawed and that their inadequacies carry a substantial cost for both the profession and the public. Where the commentators differ, both with Simon and each other, is on plausible prescriptions. While sharing Simon's commitment to the "practice of justice," the essays that follow raise substantial questions about the meaning of justice and the strategies to achieve it.

Rhode, *supra* note 53, at 867.

As Simon himself observes in responding to the contributions to the symposium:

> We have here, not the clash of opposites, but a series of family quarrels within what you might call the Party of Aspiration in legal ethics. My seven allies and I all favor a lawyers' ethic of more complex judgment and more responsibility to nonclients than the currently dominant one. The differences among us are not large from the broadest perspective, but they involve issues that are quite important to the elaboration of the sort of alternative ethic we would like to see.

William H. Simon, *The Legal and the Ethical in Legal Ethics: A Brief Rejoinder to Comments on the Practice of Justice*, 51 STAN. L. REV. 991, 991 (1999).

Indeed, differences are at times most apparent among generally similar objects or ideas. *See*

Perhaps representative of such scholars, Professor Maura Strassberg shares a number of Simon's views regarding the absence of sufficiently complex deliberation under the current model of ethical decisionmaking, but she offers an alternative significantly different from Simon's.

Like Simon, Strassberg aims to remedy the "positivist jurisprudence" that "constrains the moral content of legal ethics."[78] Tracing the evolution of the rules regulating the conduct of lawyers, Strassberg observes that "[t]here can be little doubt that the current embodiment of legal ethics in disciplinary 'codes' . . . has transformed legal ethics into positive law."[79] Yet, she concludes, "[i]n embracing positive law, the drafters of [ethics codes] did not intend to reduce ethics to positive law."[80] Rather, quoting Professor Geoffrey Hazard, Strassberg considers a code of positive law to be a "necessary preface to ethics."[81] Thus, according to Strassberg, contemporary ethics theory faces the

---

RABBI YITZCHAK HUTNER, PACHAD YITZCHAK, PURIM 42-45, 87-92 (6th ed. 1998). *Cf.* SIMON, *supra* note 1, at 248 (noting similarity between his own work and Luban's, but stating that "I have tended in public—and indeed in this book—to focus on the relatively few matters on which we disagree").

Nevertheless, some scholars offer alternative theories substantially different from Simon's approach. For example, as an alternative to Simon's view, Professor Serena Stier has proposed an "integrity thesis" aimed at "provid[ing] a better reading of the law of lawyering and hence a better interpretation of what constitutes the standard conception of lawyering." Stier, *supra* note 20, at 554. Stier provides the following summary of her proposal:

> The integrity thesis demonstrates that the correct standard conception of lawyering provides ample opportunity for attorneys to integrate their responsibilities to the profession and their own cherished moral values. Lawyers, as the Normativity Principle holds, have a special obligation to obey the law of lawyering but, as the Principle of Integrative Positivism maintains, this is a prima facie obligation only. Lawyers remain responsible for balancing their legal and moral obligations for themselves and may conscientiously decide to disobey the law of lawyering so long as their reasons for doing so are good reasons.

Id. at 609.

Applying her model to the case of the Innocent Convict, Stier concludes that there is both moral and ethical support for either silence or disclosure, and that, therefore, "integrity requires the good reader to carefully balance the respective moral and legal reasons and to act as indicated by the stronger reasons." *Id.* at 604-05. Stier offers a theory of interpreting ethics rules to permit disclosure, based on principles of the "the purpose underlying" the rules and their "underlying jurisprudence." *Id.* at 606-07. In so doing, however, her reasoning seems premised on a methodology similar to the type of "tortuous" analyses that Simon identifies among proponents of the Dominant View who attempt to find within the rules permission for disclosure to save the life of the Innocent Convict. SIMON, *supra* note 1, at 218 n.6. *See* Strassberg, *supra* note 1, at 917-18 n.93 (critiquing Stier's interpretive framework as inconsistent with Stier's defense of positivism).

78. Strassberg, *supra* note 1, at 901.

79. *Id.* at 905. *See supra* note 36.

80. Strassberg, *supra* note 1, at 910.

81. *Id.* (quoting Hazard, *Legal and Ethical*, *supra* note 36, at 12).

challenge of reconciling the positive nature of codes with a jurisprudence that avoids positivism.

In apparent contrast to Simon, however, Strassberg asserts that "given the reality of the American practice of law," it is "unrealistic" to premise a solution on either "a call for the abandonment" of current ethics rules or "a reframing of the rules."[82] Likewise, she finds unsatisfactory solutions based in "impermissible ethical disobedience,"[83] which she defines as "morally desirable action taken in apparent violation of the governing rules of conduct."[84] Requiring "reject[ion of] the existing law,"[85] such responses are "not a legal vindication" and thus "leave[] ethically disobedient lawyers with no comfort that their careers will or should survive."[86]

Instead, Strassberg insists that "[i]f we are to have legal ethics, it must be compatible with the existence of ethical rules."[87] Accordingly, she suggests, "if . . . the existing rules can be reconceptualized to make them flexible and responsive to moral concerns, then the practicing bar and disciplinary institutions may accept some degree of discretion in legal ethics."[88] Consistent with this suggestion, Strassberg offers an alternative proposal, based on the application of Ronald Dworkin's theory of "adjudication as interpretive integrity"[89] to the realm of legal ethics.[90] Thus, in Strassberg's model, "[p]ositing integrity as an ethical virtue requires the positive law of legal ethics to be read as embedded with underlying moral principles."[91]

Again similar to Simon, Strassberg illustrates her model in the context of the rules of confidentiality. Strassberg argues that application of her model to the interpretation of ethics codes would prevent the kind of "moral failure"[92] that resulted from a positivist interpretation in *Spaulding v. Zimmerman*.[93] In *Spaulding*, believing disclosure was prohibited, defense attorneys did not reveal to the plaintiff their medical expert's discovery of the plaintiff's life-threatening aneurysm.[94] According to Strassberg, "[t]he astonishing refusal of these lawyers to violate the rule may be attributed to a contemporary positivist approach to

---

82. *Id.* at 903-04.

83. *Id.* at 901.

84. *Id.* at 901 n.1.

85. *Id.* at 904.

86. *Id.* at 902-03 n.14.

87. *Id.* at 904.

88. *Id.* at 905.

89. *Id.* at 930.

90. *Id.* at 930-51.

91. *Id.* at 934.

92. *Id.* at 902.

93. 116 N.W.2d 704 (Minn. 1962). For a recent analysis of *Spaulding* and its implications, see Cramton & Knowles, *supra* note 64, at 65 & n.5-6 (noting that "*Spaulding* is extensively discussed in books and articles dealing with legal ethics and prominently featured in professional responsibility casebooks and courses," and citing sources).

94. *See Spaulding*, 116 N.W.2d 704.

legal ethics, which made disclosure in Spaulding ethically impermissible and, therefore, unthinkable."[95] Indeed, for Strassberg, "Spaulding serves as an exemplar of cases in which seemingly morally justified conduct does not fit the governing rule."[96]

Though she acknowledges that "[t]he [then-]existing language of [the rules of confidentiality] cannot provide express justification for disclosing the aneurysm to the plaintiff,"[97] Strassberg nevertheless arrives at such a result through an interpretive framework that considers the "principles arguably embedded in" the rules.[98] Specifically, Strassberg identifies a number of values underlying confidentiality, including: "[e]ncourag[ing] the full and frank client disclosure necessary for effective representation" and for "permit[ting] lawyers to effectively protect society by discouraging wrongful conduct"; "creating a 'zone of privacy' . . . , thus enhanc[ing] the autonomy and individual liberty of citizens"; and "promot[ing] the moral values of trust and loyalty."[99]

Although these values generally justify obedience to the letter of the rules of confidentiality, Strassberg asserts that in a case such as *Spaulding*, "a principle arising from a profound sense of the value of human life forces us to examine rigorously these rationales for nondisclosure."[100] Thus, she engages in an analysis that, she suggests, demonstrates the inapplicability of the rationale for confidentiality under the specific circumstances of *Spaulding*.

First, she reasons, because the information about the aneurysm was not a confidence obtained from the client, revealing the information will not affect the ordinarily salutary ramifications stemming from the open nature of attorney-client communications.[101] Second, but-for negligence on their part, the plaintiff's doctors and attorneys should have discovered the information; therefore, disclosing the information in this case is unrelated to concerns about a zone of privacy.[102] Finally, with respect to promoting trust and loyalty, "a client who caused a life and death situation betrays a moral trust by refusing to permit disclosure in order to gain financial advantage and may not be entitled to the absolute loyalty of their attorney."[103] Based on this analysis, Strassberg concludes that "disclosure of the aneurysm may be permitted" under the rules of

---

95. Strassberg, *supra* note 1, at 902.

96. *Id.*

97. *Id.* at 938.

98. *Id.* at 947.

99. *Id.* (citing ABA Comm. on Professional Ethics and Grievances, Formal Op. 250).

100. *Id. See also* Cramton & Knowles, *supra* note 64, at 87 & n.79 (noting "broad agreement" for "the primacy of human life in the hierarchy of values recognized by ordinary morality" among "[a]ll of the world's major religions" as well as "[m]oral philosophies that are secular and humanistic in nature" and concluding that "the basic proposition of respect for life is more universally accepted than perhaps any other moral tenet"). For a discussion the primacy of human life in Jewish law and tradition, see *infra* note 151.

101. Strassberg, *supra* note 1, at 947.

102. *Id.*

103. *Id.*

confidentiality.[104]

Having demonstrated that, in a case such as *Spaulding*, her interpretive model permits ethical disclosure, Strassberg turns to the broader practical effects of her proposal on the ethical deliberation of lawyers. Strassberg posits that "[l]awyers may be more likely to question whether the positive ethical law controls an apparent moral-formal dilemma when they have not been schooled to believe that moral concerns are irrelevant to interpretations of the law."[105] Moreover, she continues, if lawyers believe that "those interpreting the ethical law are not limited to the letter of the law,"[106] they will submit more requests for advisory opinions, even when "the specific language of the law appears to leave no room for avoiding the impact of the law."[107] Thus, Strassberg's model permits and encourages a greater degree of ethical discretion and deliberation on the part of lawyers, even in the face of apparently categorical rules.

II. A CRITICAL ANALYSIS OF THE SIMON AND STRASSBERG MODELS

A. Analysis of Simon's Model

Despite Simon's insistence that the Contextual View relies on current conceptions of "justice" and his observation that contextual judgment is central to modern jurisprudence,[108] his approach to confidentiality demonstrates that his solution includes, when he deems it appropriate, express rejection of contemporary ethical guidelines.[109] Indeed, Simon's use of terms such as "nullification" and "def[iance]"[110] indicates the extent to which he calls on

---

104. *Id.* at 948. *See also* Cramton & Knowles, *supra* note 64, at 87 (concluding that "[g]iven agreement about the primacy of human life as a value, the moral issue in Spaulding should be an easy one for lay people and moral philosophers alike").

105. Strassberg, *supra* note 1, at 949-50.

106. *Id.* at 950.

107. *Id.*

108. *See supra* text accompanying notes 45-49, 54-56.

109. Professor Robert Gordon has similarly observed that Simon's work can be seen as both a "fundamental" and "remarkably conservative . . . critique of the prevailing system of lawyers' ethics and practices." Robert W. Gordon, *The Radical Conservatism of The Practice of Justice*, 51 STAN. L. REV. 919, 919 (1999) [hereinafter Gordon, *Radical Conservatism*]. As Gordon explains,

> it is fundamental, in the sense that . . . Simon razes to the ground the current structure of ethical rules and their presuppositions. It is conservative, in that he then shows how a system of lawyers' ethics can be rebuilt on its existing foundations, using existing construction materials—the ordinary working conceptions of law and justice that lawyers bring to bear in other aspects of their practices.

Id. Cf. Robert W. Gordon, *Portrait of a Profession in Paralysis*, 54 STAN. L. REV. 1427, 1427 (2002) (reviewing RHODE, *supra* note 43, and stating that Rhode "has many practical concerns for reform, which on the whole are modest, incremental, sensible, and entirely feasible").

110. SIMON, *supra* note 1, at 164.

lawyers consciously to violate ethics rules.[111] To be sure, Simon is not alone among ethics scholars who prescribe violation of the rules of confidentiality to save the life of the Innocent Convict.[112] In fact, there is nothing particularly remarkable about such a position, when articulated as a moral judgment based in considerations of the competing values of confidentiality and saving a life.[113] Yet, Simon's approach is different, in that he refuses to characterize his solution as contrary to law; instead, Simon refers to violation of the rules as "an act of principled commitment to legal values more fundamental than those that support the rule."[114] Notwithstanding his own characterization of the result, however, Simon's solution is ultimately premised on a rejection of and consequent disregard for prevailing ethics regulations.[115]

---

111. Likewise, Simon acknowledges, albeit in a rather understated fashion, the clear nature of the legal authority that supported the prohibition against disclosure of confidential information in the case of the Innocent Convict, stating that "many people, *including people with authority over the matter*, accept . . . the arguments for categorical confidentiality safeguards." *Id.* (emphasis added). Moreover, he notes that: "[f]or the most part the norms have been recently enacted, often after substantial debate"; "the bar's rules are consistent in important respects with the attorney-client privilege in the law of evidence"; and "professional responsibility norms are ultimately enacted by courts." *Id.* Nevertheless, in the face of this demonstration of the legal legitimacy of the rule prohibiting disclosure, Simon concludes that these "considerations" are "outweighed" by the arguments he offers in favor of disclosure. *Id.*

112. *See, e.g.*, Cochran, *supra* note 75; Mary C. Daly, *To Betray Once? To Betray Twice?: Reflections on Confidentiality, A Guilty Client, an Innocent Condemned Man, and an Ethics-Seeking Defense Counsel*, 29 LOY. L.A. L. REV. 1611 (1996); Robert P. Lawry, *Damned and Damnable: A Lawyer's Moral Duties With Life on the Line*, 29 LOY. L.A. L. REV. 1641 (1996); Russell G. Pearce, *To Save a Jewish Life: Why a Rabbi and a Jewish Lawyer Must Disclose a Client Confidence*, 29 LOY. L.A. L. REV. 177 (1996).

113. *See, e.g.*, Cochran, *supra* note 75, at 577 ("I think that the moral arguments for revealing [the convict's] innocence, the importance of human life, and protecting the innocent, justify a lawyer revealing this information . . .); Pearce, *supra* note 112, at 1776 (quoting *Leviticus* 19:16; *Deuteronomy* 30:19) (concluding that, on the basis of Jewish law, the lawyer should reveal the information, because "the combined duties 'not to stand idly by' and 'to preserve life' outweigh [ ] the Jewish legal obligations of confidentiality" and following the law of the land).

114. SIMON, *supra* note 1, at 164.

115. *See* Cochran, *supra* note 75, at 579-80 (observing that "Simon insists that he is making a legal judgment" in favor of disclosure, but concluding that Simon's analysis is "short on legal support" and noting that "Simon cites no legal sources that justify revealing the confidential information"); Feldman, *supra* note 53, at 1479 (observing that Simon "devotes a large part of the book to justifying nullification, which he argues is not really nullification but, when done rightly, is correction of a misunderstanding of what the law is"); Gordon, *Radical Conservatism*, *supra* note 109, at 922 (stating that Simon's "view of 'law' is one of judgments that often, though not invariably, incorporate moral norms, including norms that sometimes justify ad hoc nullification or even conscientious resistance to laws whose operation is conspicuously unjust"); Luban, *supra* note 53, at 886 (challenging Simon's premise that "any argument for disobedience against a particular command would also be an argument that the command was an incorrect interpretation

of the law"); Strassberg, *supra* note 1, at 904 (observing that Simon's model "appears incompatible with the current rules of legal ethics" and finding it "unlikely . . . that the organized bar would be sympathetic to [Simon's] reframing of the rules").

See also Cochran, *supra* note 75, at 576 (acknowledging that "[t]he law of professional responsibility can be a source of injustice" and that, in the case of Leo Frank, the actual case on which Simon modeled his discussion of the Innocent Convict, "[t]he law governing lawyers led to the most unjust of results"); Cramton & Knowles, *supra* note 64, at 65 (observing "the reality . . . that the adversary role of the lawyer in litigation arguably permits, and may sometimes require, a lawyer to behave in an amoral or immoral way"); Luban, *supra* note 53, at 887 (recognizing that "some laws are morally unacceptable under any interpretation that does not do violence to the text" and that "[e]ven good laws occasionally yield unjust outcomes in exceptional cases").

Simon's approach is thus significantly different from the position of others who likewise prescribed disclosure to save the life of the Innocent Convict but acknowledged that their solution was inconsistent with the legal obligations of lawyers. *See* Cochran, *supra* note 75, at 577-79 n.39 (stating that "'legal values' appear to be on the side of not revealing" the information, but concluding that "I agree with Simon that the lawyer should reveal the information, but as a matter of moral judgment, rather than legal judgment"); Daly, *supra* note 112, at 1612-13 (concluding that "I am confident of my response" that "[m]y conscience will not permit confidentiality to trump justice," because "[a]n innocent man cannot be allowed to die"). Daly describes it as "unambiguous and consistent," that whether the issue is

> filtered through the prism of the attorney-client privilege or that of the ethical obligation of confidentiality; whether it is analyzed historically under the Canons or the Model Code; or whether it is examined newly under the Model Rules[,] [t]he answer is the same: [the lawyer] may not disclose [the confidence] even to save the life of an innocent man.

Id. at 1621-22; Lawry, *supra* note 112, at 1652, 1654 (identifying "a moral imperative" to disclose the information and stating that "[i]f the state or the legal profession punished me for disclosing the information . . . then so be it"); Pearce, *supra* note 112, at 1774, 1779 (stating that, under the rules, "one could assume that confidentiality as a lawyer is required" but concluding that the lawyer "had to do whatever she could to try to save a life," even though she would be "jeopardizing her career as a lawyer"); Rosenthal, *supra* note 16, at 166, 168 (stating that "[a]ttorneys who decide to disclose under the circumstances [of the Innocent Convict] may maintain a sense of moral equilibrium by adhering to their principles" but they "must be prepared for the potential consequences" of violating the rule).

See also John Leubsdorf, *Using Legal Ethics to Screw Your Enemies and Clients*, 11 GEO. J. LEGAL ETHICS 831, 834 (1998) ("Sometimes the law requires lawyers to act unethically. Blame then falls on the law-makers, and the lawyer's only options are civil disobedience and the pursuit of law reform."); Luban, *supra* note 53, at 889 (recognizing that at times, "the lawyer confronts a law-versus-morality question (not a law-versus-law question): 'Should I commit a crime to prevent injustice?'"); Stier, *supra* note 20, at 607-08 n.233 ("[T]he essence of civil disobedience [is] that it may be morally required although it both demands exceptional courage to undertake and resolution to endure its penalties").

Rosenthal notes the possibility of discipline resulting from disclosure, though he emphasizes that "[f]rom an economic standpoint, a more tangible, often damaging consequence exists" because

Moreover, and more significantly for the purposes of the present analysis, despite his declared goal of incorporating contextual considerations into the ethical judgment of attorneys, Simon's model does not appear to increase the ethical discretion of individual lawyers. Rather, Simon's approach seems to replace arguably unsatisfactory dictates of the current regulatory model with alternative—yet similarly mandatory—prescriptions in response to complex ethical questions. Simon's solution to the problem of the Innocent Convict is again instructive.

Simon's discussion of the Innocent Convict is certainly compelling. There is undoubtedly substantial appeal to his position that saving the life of the Innocent Convict outweighs any considerations supporting a categorical rule of confidentiality for information related to completed crimes. Even among scholars who disagree with Simon's characterization of his reasoning as based in legal—in addition to moral—principles, there is much agreement with the substance of his analysis.[116]

Nevertheless, it would seem presumptuous to assume that disclosure is the only viable ethical response. Clearly, both the organized bar, as reflected in the Model Code and the Model Rules, and many state courts, as reflected in the implementation of similar rules, adopted the position that, despite the tragic consequences to the Innocent Convict, a lawyer's ethical duty requires maintaining client confidences.[117] Though this position was the subject of much

---

"[t]he reputation of criminal defense attorneys travels swiftly through the ranks of criminal defendants and once the attorney is labeled as untrustworthy, that attorney may likely be hard pressed to retain any future clients." Rosenthal, *supra* note 16, at 167-68. Therefore, he concludes, "the decision to disclose . . . in this scenario is analogous to professional martyrdom." *Id.* at 168. *Cf.* Stier, *supra* note 20, at 598 n.185 ("It would be both self-righteous and foolish to prescribe martyrdom as an essential ingredient of integrity").

116. *See supra* note 115.

117. For a description—and sharp critique—of the initial refusal of the American Law Institute ("ALI") to provide an exception to the rules of confidentiality in this scenario, see Monroe H. Freedman, *The Life-Saving Exception to Confidentiality: Restating the Law Without the Was, the Will Be, or the Ought to Be*, 29 LOY. L.A. L. REV. 1631 (1996). Professor Freedman refers to this refusal as a "bizarre . . . example of th[e] tendency to overlook the ought [to be]," *id.* at 1631, and states that "[t]he fact that the questions posed in these cases are seriously debated within the legal profession is itself a kind of sick lawyer joke." *Id.* at 1632.

Simon is likewise critical of the approach of the ALI, observing in particular that a draft of the Restatement had included the case of the Innocent Convict as an illustration of the rule prohibiting disclosure in such a scenario. As Simon notes, with open incredulity, "[a]fter heated debate, the members [of the ALI] voted to delete the illustration as 'offensive' *without making any change to the rule!*" SIMON, *supra* note 1, at 222 n.9 (emphasis in original) (quoting 5 LAWYER'S MANUAL ON PROFESSIONAL CONDUCT (ABA/BNA) 1581-59 (1989)).

It should be noted that, like the current version of the Model Rules, the final version of the Restatement permits disclosure "to prevent reasonably certain death or bodily harm to a person." RESTATEMENT, *supra* note 36, § 66. The illustration accompanying the rule is taken from the facts of *Spaulding*, though the rule would now permit disclosure in the case of the Innocent Convict as

controversy and criticism, it was based on the conclusion of these groups and authorities that on some level and in some cases, contrary to Simon's reasoning, the benefits and values underlying nearly categorical rules of confidentiality outweigh saving a human life.[118]

---

well. *See id.* at § 66, illus. 1.

118. *See, e.g.*, Kathryn W. Tate, *The Hypothetical as a Tool for Teaching the Lawyer's Duty of Confidentiality*, 29 LOY. L.A. L. REV. 1659, 1683 (1996) (arguing that "[a]llowing or mandating attorneys to reveal a client's confidences puts the profession on a slippery slope of having to be the judge of which confidences are to be revealed and which are not" and concluding that "[n]ow that the profession has begun to assume the role of deciding when to make disclosures in the first instance, without judicial scrutiny and without clear guidelines, lawyers may become more and more comfortable with this role change and with disclosure," and as a result, "the client will only have the luck of the draw as to what moral standard or perspective of the professional code's scope the lawyer selected will follow").

Professor Abbe Smith advocates "a strict principle that client secrets and confidences are sacrosanct and lawyers should not divulge them under any circumstances," arguing that "it is more important to maintain and preserve the principle of confidentiality—no matter how difficult the circumstance—than it is to affirm individual lawyer morality." FREEDMAN & SMITH, *supra* note 42, at 147. Smith is also concerned that under more permissive rules, "lawyers will be more likely to use or disclose a client's confidential information . . . when the client is an indigent . . . than any other client," resulting in "an even greater divide between the kind of legal services—and loyalty—provided to some clients than that provided to others." *Id.* Moreover, Smith "worries . . . that there will be a spillover from lawyers who too broadly disclose client information in the name of third parties." *Id.* According to Smith, "[t]here is already a prevalent view that court-appointed lawyers are not to be trusted, that they don't care about their clients." *Id.* She concludes that "[p]oor clients will soon stop disclosing information to lawyers altogether—and perhaps they would be wise to do so." *Id.*

It should be noted that Smith considers the scenario of the Innocent Convict a "rare case where it is truly necessary to disclose information obtained through the lawyer-client relationship." *Id.* She anticipates that under such circumstances, notwithstanding a strict rule of confidentiality, a lawyer would disclose the information but would not be disciplined. *Id.*

Cf. Cramton & Knowles, *supra* note 64, at 112-13 (concluding that "a person's interest in averting wrongful execution or incarceration justifies disclosure" but "recogniz[ing] that client betrayal is likely to be more troublesome in situations in which the disclosure may result in the client being punished for the crime for which another person has been wrongfully convicted"); Lawry, *supra* note 112, at 1652-53 (concluding that disclosure is "a moral imperative" but stating that "such a breach would not be easy" in light of justifications for general confidentiality).

Professor Stier has argued that, in the case of the Innocent Convict, "there are good moral and legal reasons available to support either silence or disclosure." Stier, *supra* note 20, at 605. Specifically, "[p]reserving client confidences because of the promise of loyalty is a good moral reason for maintaining one's silence[, and s]uch an action can readily find legal support in the confidentiality provisions of the professional standards." *Id.* However, "[u]nder the facts of this unusual and extremely troubling hypothetical, revealing client confidences in order to save an innocent life obviously can be justified morally." *Id.* at 605-06. Moreover, according to Stier, "[s]upport for this decision can also be located in the standards of the profession, but less readily."

According to Simon, however, such a conclusion is not merely problematic, or even disturbing; it is plainly wrong, both in its outcome and its rationale. In Simon's view, the only correct analysis in the case of the Innocent Convict is one that, on balance, considers saving a life more important than any justifications for maintaining confidences. Therefore, it follows, in Simon's view, the only correct outcome in the case of the Innocent Convict is the one he prescribes: disclosure of confidences to save a life. Perhaps Simon's solution offers a more complex mode of analysis than that suggested by the rules, in the sense that it considers more factors, or at least engages and then attempts to refute arguments supporting the rule. Nevertheless, in concluding that disclosure is the only viable ethical solution, Simon likewise prescribes an apparently categorical answer to a complex ethical question, resulting in a form of ethical decisionmaking and adherence seemingly similar to the mechanical method he so strenuously rejects.[119]

B. Analysis of Strassberg's Model

Similarly, an examination of Strassberg's position suggests that, despite her efforts to the contrary, like Simon's, her solution appears to raise concerns over both its viability within prevailing ethical standards and, more importantly, its ability to effect the broader aim of increased ethical deliberation. First, despite Strassberg's stated goal of working within "the interpretation and application of the existing rules of ethics, as set out in either the Model Code or the Model Rules,"[120] her solution appears vulnerable to criticism similar to that which she levels against Simon and others.[121] Second, although Strassberg's model offers a broader range of ethical discretion for lawyers than do categorical rules, again, like Simon's approach, it appears to have a limited effect in the extent to which it will result in more careful and complex ethical deliberation by lawyers.

Although her analysis offers a method for interpreting current rules regulating lawyers, her proposal seems unrealistic in light of—if not ultimately inconsistent with—the substance of the rules. Strassberg insists that, unlike Simon and others, whose work she characterizes as calling for an "abandonment"

Id. at 606.

119. *See* Wendel, *supra* note 5, at 73 (stating that "[o]ne might ask whether Simon places more weight on the objectivity of legal judgments than is warranted" and that Simon "sounds in places like [Ronald] Dworkin, who has sometimes argued that all legal questions have a single right answer"); *cf.* Luban, *supra* note 53, at 895 (describing Simon as "Dworkinian, in that the interpretation of legal standards turns out to presuppose background views about difficult questions of political theory into which the lawyer must weave the legal standard in the most coherent way possible"). *But see* Atkinson, *supra* note 75, at 1517 (concluding that Simon's book "is thus best read . . . not as a blueprint for replacing one professional order of mandatory rules and sanctions with another more to his liking, but as an invitation to join him in examining our shared professional lives").

120. Strassberg, *supra* note 1, at 923.

121. *See supra* text accompanying notes 82-86.

or a "reframing"[122] of ethics rules, her own model involves a "reconceptualiz[ation]" of the rules.[123] Notwithstanding the theoretical distinction between these characterizations, and even looking beyond positivist jurisprudence, ultimately, Strassberg's conclusion—that a rule expressly and categorically prohibiting disclosure of confidences may be understood to permit such disclosure—seems more a rejection of the rule than a reinterpretation.

In addition, whereas Simon's model may prescribe an assertedly correct answer to an ethical dilemma, thereby obviating the need for ethical deliberation on the part of the individual lawyer,[124] Strassberg's solution seems merely to allow, but not require, that the lawyer engage in ethical deliberation. Strassberg does encourage lawyers to interpret ethics rules in light of moral principles, thus supporting a more complex interpretation of ethical guidelines. For example, Strassberg offers an interpretive framework through which facially categorical confidentiality rules may be understood, after careful analysis, to permit disclosure in cases such as *Spaulding*.[125] Nevertheless, Strassberg's model does not appear to impose on lawyers the obligation to engage in such analysis. Although some lawyers will perhaps accept the challenge of applying more nuanced ethical considerations to the interpretation and implementation of ethics codes, Strassberg's solution does not seem to suggest a way of increasing or improving the level of ethical analysis among those who continue to opt for more mechanical adherence to ethical guidelines.

C. Analysis Under Amended Model Rule 1.6

Addressing the concerns underlying the broad criticism of the Model Rules' facial prohibition of disclosure in scenarios such as the Innocent Convict and *Spaulding*, in 2002, the American Bar Association adopted an important change to Model Rule 1.6. The amended rule allows disclosure "to prevent reasonably certain death or substantial bodily harm."[126] Thus, amended Model Rule 1.6 apparently permits a lawyer in the Innocent Convict scenario to reveal that the client, not the Innocent Convict on death row, committed the capital crime, or in the Spaulding case to reveal an adversary's life-threatening injury caused by the client.[127]

---

122. Strassberg, *supra* note 1, at 904.

123. *Id.* at 905.

124. *See supra* Part II.A.

125. *See supra* Part I.B.

126. MODEL RULES, *supra* note 34, R. 1.6 (b)(1). *See also* RESTATEMENT, *supra* note 36, § 66 ("A lawyer may use or disclose confidential client information when the lawyer reasonably believes that its use or disclosure is necessary to prevent reasonably certain death or serious bodily harm to a person.").

127. The amended comment to the rule incorporates the rationale of those who have advocated disclosure in these scenarios, emphasizing that "[a]lthough the public interest is usually best served by a strict rule requiring lawyers to preserve the confidentiality of information relating to the representation of their clients," the amended rule "recognizes the overriding value of life and

On one level, the amended rule appears to alleviate the concerns of Simon, Strassberg, and others, proposing expressly incorporating into the law governing

---

physical integrity . . . ." MODEL RULES, *supra* note 34, R. 1.6, cmt. [6]. Although the comment does not refer to either the Innocent Convict or *Spaulding*, it presents a similar—perhaps more common—scenario, in which:

> a lawyer who knows that a client has accidentally discharged toxic waste into a town's water supply may reveal this information to the authorities if there is a present and substantial risk that a person who drinks the water will contract a life-threatening or debilitating disease and the lawyer's disclosure is necessary to eliminate the threat or reduce the number of victims.

Id.

In adopting this amendment to Rule 1.6, the ABA accepted only one component of a considerably more extensive amendment to the rule proposed by the Ethics 2000 Commission. The proposal included two additional provisions permitting disclosure, in limited circumstances, to prevent financial harm to a third party. After the first of these provisions was rejected by the ABA in 2001, the other was withdrawn. *See* GILLERS & SIMON, *supra* note 63, at 80-81.

These proposals and the decisions of the ABA are, to some degree, representative of the broader work of the Ethics 2000 Commission and the responses of the ABA. For example, the proposed amendments to Model Rule 1.6 appeared to be an attempt by the Ethics 2000 Commission to reflect the judgments and values underlying proposals that had likewise been considered but rejected before the initial enactment of the Model Rules in 1983—though it should be noted that the earlier proposals included mandatory disclosure provisions as well. *See id.* at 80-81. Similarly, the Ethics 2000 Commission proposed an amendment to Model Rule 1.5 that would have revived earlier drafts of the rule requiring that fee agreements be put in writing, in place of the enacted rule's language merely suggesting that fee agreements should "*preferably* [be] in writing." *See id.* at 58-59; MODEL RULES, *supra* note 34, R. 1.5(b) (emphasis added). The ABA rejected this proposal as well. *See* GILLERS & SIMON, *supra* note 63, at 59.

On the other hand, the Ethics 2000 Commission did achieve an arguably substantial level of success with respect to Model Rule 1.6, effecting a change in the confidentiality rule to provide an exception to prevent death or substantial physical harm. Another example of a successful proposal, adopted by the ABA, requires that a lawyer entering into a business transaction with a client not only give the client "a reasonable opportunity" to seek the advice of a lawyer regarding the transaction, but also advise the client, in writing, of "the desirability" of seeking such counsel. MODEL RULES, *supra* note 34, R. 1.8.

Indeed, not surprisingly, as with most legislative processes—including the drafting of ethics codes—the responses of the ABA to the proposals of the Ethics 2000 Commission have been varied and complex. *See* Margaret Colgate Love, *The Revised ABA Model Rules of Professional Conduct: Summary of the Work of Ethics 2000*, 15 GEO. J. LEGAL ETHICS 441 (2002); *cf.* Richard Abel, *Why Does the ABA Promulgate Ethical Rules?*, 59 TEX. L. REV. 639 (1981); Marvin Frankel, *Why Does Professor Abel Work at a Useless Task?*, 59 TEX. L. REV. 723 (1981); Stephen Gillers, *What We Talked About When We Talked About Ethics: A Critical View of the Model Rules*, 46 OHIO ST. L.J. 243 (1985); Thomas Morgan, *The Evolving Concept of Professional Responsibility*, 90 HARV. L. REV. 702 (1977); Deborah Rhode, *Why the ABA Bothers: A Functional Perspective on Professional Codes*, 59 TEX. L. REV. 689 (1981); Ted Schneyer, *Professionalism as Bar Politics: The Making of the Model Rules of Professional Conduct*, 14 LAW & SOC. INQUIRY 677 (1989).

lawyers the moral principles recognizing the overriding value of human life. As a consequence, the letter of the amended rule permits the very discretion to disclose that Simon and Strassberg struggled to justify, thereby obviating the need for theoretical interpretations and reconceptualizations that seem inconsistent with the prior formulation of the rule.

Nevertheless, a closer look at the amended rule suggests that it neither yields the same result as, nor meets the goals of Simon's model. In his discussion of the case of the Innocent Convict, after weighing the competing ethical considerations, Simon prescribes disclosure as the correct decision, in apparent violation of the former rule.[128] Undoubtedly, then, the permission to disclose provided under the current rule would, according to Simon, imply an obligation to disclose. Indeed, Simon hypothesizes that "[i]f . . . the rules gave the lawyer discretion to disclose, then the case would be an easy one."[129]

In contrast to Simon's hypothetical analysis, however, the current rule, despite permitting disclosure, does not suggest that the case is an easy one militating an obligation to disclose. Rather, notwithstanding the comment to the rule describing the "overriding value of life,"[130] the rule does not definitively determine that saving a life overrides confidentiality; instead, it only permits such a discretionary conclusion on the part of the lawyer. Hence, the current rule differs significantly from Simon's model in that it continues to allow a lawyer, alternatively, to conclude—contrary to Simon's view of ethics and morality—that confidentiality outweighs saving the life of the Innocent Convict. Thus, unlike Simon's model, the current rule arguably expands the range of ethical discretion for the individual lawyer, who ultimately decides among competing values.

The current rule appears more consistent with the outcome of the *Spaulding* scenario under Strassberg's model, though the current rule may also be susceptible to the same limitations as Strassberg's model. Like Simon, despite operating under the prior version of Rule 1.6, Strassberg presented a theory permitting disclosure to save a life. Yet, unlike Simon, Strassberg did not conclude that disclosure is mandatory in such a case.[131] Therefore, even under the more permissive current rule, although the case for disclosure may seem more compelling, Strassberg would appear to permit—but still not require—disclosure. Thus, both Strassberg's model and the current rule may be intended to increase the likelihood as well as the quality of ethical deliberation. In practice, however, although they may encourage such analysis on the part of lawyers, they do not offer a framework through which such deliberation becomes obligatory.

As an alternative to both Simon's and Strassberg's models, perhaps a more effective method for increasing and improving ethical deliberation among lawyers requires a different understanding of the concept of discretion in legal ethics.

---

128. *See supra* Part II.A.

129. SIMON, *supra* note 1, at 163.

130. MODEL RULES, *supra* note 34, R. 1.6, cmt. [6].

131. *See supra* Part II.B.

III. BEYOND THE MANDATORY/OPTIONAL DICHOTOMY: A DELIBERATIVE MODEL OF LEGAL ETHICS

Under the standard conception of ethics rules, an ethics provision is either mandatory or optional. This conception is reinforced, in part, by the "Scope" section introducing the Model Rules, which states that "[s]ome of the Rules are imperatives, cast in the terms of 'shall' or 'shall not.'"[132] These Rules "define proper conduct for purpose of professional discipline."[133] In contrast, "[o]ther [Rules], generally cast in the term 'may,' are permissive and define areas under the Rules in which the lawyer has discretion."[134] With respect to the latter, "[n]o disciplinary action should be taken when the lawyer chooses not to act or acts within the bounds of such discretion."[135] Thus, it appears that if an ethics provision does not contain an express command, in the form of required or prohibited conduct, the lawyer may decide whether and in what manner to respond to the provision, and may not be disciplined for this decision.[136]

Yet, there may be an alternative conceptual model for ethics rules, a "Deliberative Model," that looks beyond a simplistic mandatory/optional dichotomy. Under this model, in addition to the categories of rules that are mandatory and optional, there is a third category of "discretionary" rules. Significantly, in this conceptual framework, a discretionary rule differs from an optional rule. Adherence to a discretionary rule, as the term implies, is—like adherence to an optional rule—ultimately at the discretion of the lawyer. However, unlike the suggestions set forth in optional rules—consideration of which is in no way obligatory—the ethical alternatives presented in a discretionary rule must be considered and deliberated before a decision is made.[137] Thus, whereas an optional rule is purely suggestive—often

---

132. MODEL RULES, *supra* note 34, Scope [14].

133. *Id.*

134. *Id.*

135. *Id.*

136. It should be noted that even under this understanding of ethics rules, ethics provisions need not always state, with complete specificity, the precise nature of conduct that may be subject to discipline. Both the Model Code and the Model Rules contain broad ethics provisions that courts have interpreted as a basis for discipline, finding violations of what I have called elsewhere "unenumerated ethical obligations." *See* Levine, *supra* note 36; *see, e.g.*, MODEL CODE, *supra* note 34, DR 1-102 (A)(6) (prohibiting "conduct that adversely reflects on [the lawyer's] fitness to practice law"); MODEL CODE, *supra* note 34, DR 1-102 (A)(5) (prohibiting "conduct that is prejudicial to the administration of justice"); MODEL RULES, *supra* note 34, R. 8.4(d) (same); MODEL CODE, *supra* note 34, Canon 9 ("avoid[ing] even the appearance of professional impropriety"). For examples of specific conduct that courts have held to be in violation of these broad ethics provisions, see Levine, *supra* note 36.

137. The Deliberative Model borrows in part from Rabbi Yitzchak Hutner's analysis delineating the nature of different activities under Jewish law and philosophy. *See* YITZCHAK HUTNER, PACHAD YITZCHAK, PESACH 123-26 (6th ed. 1999). Rabbi Hutner refutes a misconception

aspirational[138]—a discretionary rule must be taken seriously before a lawyer

---

that categorizes a limited number of human actions as either obligatory or prohibited pursuant to God's express or implied command, but considers purely optional the remainder of life's activities. *See id.* at 123-24.

As Rabbi Hutner explains, the range of activities that are truly optional, in the sense of being exempt from moral consideration, is virtually nonexistent. Instead, based on the biblical verse, "in all of your ways you shall acknowledge [God]," *Proverbs* 3:6, Rabbi Hutner describes a broad and general obligation that all of life's activities be used to serve God. However, he continues, unlike those activities that are expressly or implicitly identified as mandated or prohibited under Jewish law, the remaining activities require a more complex analysis to be exercised, under a consideration of the relevant circumstances, in the service of God. *See id.*; *see also* HUTNER, *supra* note 77, at 51-53; RABBENU BACHYA IBN PAQUDA, CHOVOTH HA-LEVAVOTH (4:4); Levine, *Broad Life*, *supra* note 7; Levine, *Professionalism*, *supra* note 7; Samuel J. Levine, *Unenumerated Constitutional Rights and Unenumerated Biblical Obligations: A Preliminary Study in Comparative Hermeneutics*, 15 CONST. COMMENT. 511 (1988); Aharon Lichtenstein, *Does Jewish Tradition Recognize an Ethic Independent of Halakha?*, CONTEMPORARY JEWISH ETHICS 102-23 (Menachem Marc Kellner ed., 1978); Moshe Sokol, *Personal Autonomy and Religious Authority*, *in* RABBINIC AUTHORITY AND PERSONAL AUTONOMY 169-216 (Moshe Sokol ed., 1992).

Similarly, the Deliberative Model observes that some ethics rules specify actions that are mandated or prohibited, thus determining definitively the ethical nature of these actions. However, the fact that other rules identify actions that are left to the ethical discretion of lawyers does not imply that such conduct is optional in the sense that it is outside of the realm of ethical consideration. Instead, the rules acknowledge that such conduct is not easily or universally defined as ethically mandated or prohibited. Thus, under these rules, the lawyer is obligated to engage in ethical deliberation to determine the way in which, under the circumstances, the rule can best be applied in an ethical manner.

As I have elsewhere noted and developed at greater length, hermeneutic comparison between Jewish law and American legal ethics may be helpful in part "because both focus on and prescribe obligations rather than rights." Levine, *supra* note 36, at 543. Likewise, a comparison of the modes of ethical decisionmaking in the two systems may be particularly appropriate, as both involve consideration and deliberation of complex legal and moral issues underlying these obligations.

138. *See, e.g.*, MODEL RULES, *supra* note 34, R. 6.1 ("A lawyer should aspire to render at least (50) hours of pro bono publico legal services per year."). The clearly aspirational nature of the Rule is demonstrated by the use of the word "should," which indicates an optional suggestion, rather than a statement of either mandatory, prohibited, or discretionary conduct. In fact, the Scope section likewise distinguishes between: (1) "Rules [that] are imperatives, cast in the terms 'shall' or 'shall not'"; (2) Rules, "cast in the term 'may,' [that] are permissive and define the areas under the Rules in which the lawyer has discretion to exercise professional judgment"; and (3) Comments that use the term "should" and do not add obligations to the Rules. MODEL RULES, *supra* note 34, Scope Cmt. [14]. Thus, consistent with the distinctions described in the Scope, the pro bono rule, employing the term "should," does not impose additional obligations, while discretionary rules, which use the term "may," require the exercise of professional judgment.

The Deliberative Model thus finds much in common with, but ultimately differs from, Professor Russell Pearce's proposed Rule that "would hold lawyers morally accountable for their

decides to act or not to act.[139]

---

conduct." Pearce, *supra* note 25, at 1805. Like the Deliberative Model, Pearce's proposal "would not dictate a particular moral vision. Rather, it would direct lawyers with sometimes conflicting understandings of their role to wrestle with the moral implications of their conduct both as individuals and as a community." *Id.* The Deliberative Model, likewise, would not dictate a particular view of a complex moral issue, but would instead impose the obligation to engage in ethical deliberation before determining a course of action.

Unlike the Deliberative Model, however, Pearce describes the proposed Rule 1.0 as "aspirational, similar to Rule 6.1." *Id.* at 1807. Thus, "[t]he goal of the Rule would be to educate lawyers to their moral responsibility and to encourage lawyers as individuals and as members of the legal community to explore how their work 'contribut[es] to the social good.'" *Id.* (quoting RHODE, *supra* note 43, at 8). In contrast to both Model Rule 6.1 and Pearce's proposed Model Rule 1.0, the Deliberative Model envisions an obligation, not merely an option or aspiration, of ethical deliberation and decisionmaking.

139. The lack of precision in the current approach to discretionary ethics rules may have contributed to the confusion that has arisen in the interpretation and application of provisions stating that an attorney "may" reveal confidences when "required" by a court order. For example, a case before the Supreme Court of Illinois involved a rule stating that a lawyer "*may*" reveal "'confidences or secrets when . . . required by law or court order.'" In re Marriage of Decker, 606 N.E.2d 1094, 1103 (Ill. 1992) (emphasis in original) (quoting 107 Ill.2d R. 4-101). The Illinois State Bar Association argued that the rule was "discretionary" because it used the "permissive word 'may'" and not the "mandatory 'shall,'" and that, therefore, under the rule, an attorney was not required to disclose confidences pursuant to a court order. *Id.* The court rejected this interpretation, concluding that the phrase "required by law or court order" is "clearly mandatory." *Id.* (quoting Ill.2d R. 101(d)(2)). Accordingly, the court explained that "rather than being discretionary, [the rule] simply instructs attorneys that they will not be disciplined for revealing confidences or secrets when *required* by law or court order." *Id.* at 1103-04 (emphasis in original). The court added that "[t]o hold otherwise would place an attorney's discretion above judicial determination of the matter." *Id.* at 1103. *See also* Fellerman v. Bradley, 493 A.2d 1239, 1248 (N.J. 1985) (interpreting a similar rule and concluding that "the attorney must comply with the trial court's order compelling disclosure.").

The differing interpretations in these cases may be an unavoidable consequence of the problematic wording of the rules. Alternatively, the cases may serve as illustrations of Professor Koniak's view that "[t]he state and the profession have different understandings of the law governing lawyers—they have in effect different 'law,'" as well as examples of what Professor Koniak sees as a "continuing struggle between the profession and the state over whether the profession's vision of law or the state's will reign." Koniak, *supra* note 60, at 1390.

Nevertheless, it appears that some of the confusion in the interpretation of these rules stemmed from imprecise use and application of the concepts "mandatory," "discretionary" and "permissive" as applied to ethics regulations. The Illinois Bar Association appears to have premised its argument upon the understanding that the rule was discretionary and, therefore, optional. This understanding was based in part on a parallel interpretation of Rule 4-101(d)(3), which provided that an attorney may reveal the intention of a client to commit a crime that would not involve death or serious injury. As the court acknowledged, this rule "is discretionary" and therefore, under this rule, the lawyer "had the discretion to inform" or not to inform. *In re Marriage of Decker*, 606 N.E.2d at

Thus, whereas mandatory rules prescribe a particular mode of conduct, violation of which may subject a lawyer to discipline, discretionary rules under the Deliberative Model prescribe ethical deliberation.[140] Therefore, if a response to a discretionary rule is determined on the basis of an articulable and justifiable form of ethical deliberation, regardless of the outcome, discipline is not appropriate. If, however, a course of conduct is undertaken without such deliberation, depending on the circumstances, such conduct may subject the lawyer to discipline.[141]

---

317. Because the court and the Bar Association both equated discretionary rules with optional rules, the Bar Association simply applied the same interpretation to Rule 4-101(d)(2). In response, the court could distinguish Rule 4-101(d)(2) only by asserting that it was not a discretionary rule.

Perhaps a more helpful understanding of these rules would view both of them as discretionary rules, in the sense of requiring the lawyer's ethical deliberation in their interpretation and application. Under such an analysis, Rule 4-101(d)(3) would continue to allow for either disclosure or silence on the basis of ethical deliberation. Consistent with the court's conclusion, however, ethical consideration of Rule 4-101(d)(2) would require disclosure pursuant to a court order, because the lawyer's ethical discretion, quite understandably, would not override the implicit judicial determination of the ethical issue through its order to disclose.

140. The precise wording of the "Scope" section lends some support to this conception. The sentence stating that permissive rules "define areas . . . in which the lawyer has discretion" concludes with the words "to exercise professional judgment." MODEL RULES, *supra* note 34, Scope Cmt. [14]. This language seems to indicate that a lawyer's discretion in relation to a permissive rule does not include an option blithely to disregard the ethical concerns underlying the rule. Rather, with discretion comes the responsibility to exercise professional judgment. Thus, although the lawyer ultimately may decide how to respond to a permissive rule, such a decision is acceptable only when reached as the result of exercising professional judgment, presumably through ethical deliberation. Simon has similarly emphasized the centrality of the concept of professional judgment, both in the area of legal ethics and more generally in the American legal system. *See* SIMON, *supra* note 1, at 198-99. Because discretion is thereby based in professional judgment, it is arguably only in the exercise of such judgment that, as the Scope continues, "[n]o disciplinary action should be taken when the lawyer chooses not to act or acts within the bounds of such discretion." MODEL RULES, *supra* note 34, Scope Cmt. [14]. The complex and limited nature of the obligation imposed by discretionary rules may be captured in the seemingly oxymoronic reference to "permissive duties." Fred C. Zacharias, *Five Lessons for Practicing Law in the Interests of Justice*, 70 FORDHAM L. REV. 1939, 1943 (2002).

141. Scholars have documented the extent to which, on a descriptive level, the substance of many ethics rules consciously "do not claim to occupy the whole field of ethical deliberation." Wendel, *supra* note 5, at 11. Professor Wendel delineates two different categories of such rules: First, "[s]ome rules expressly build in an element of discretion, such as the provision which empowers a lawyer to withdraw from representation if 'a client insists upon pursuing an objective that the lawyer considers repugnant,'" *id.* (quoting MODEL RULES, *supra* note 34, R. 1.16(b)(3)) (presently R. 1.16 (b) (4)), "or the rule which permits a lawyers to charge only reasonable fees." *Id.* (quoting MODEL RULES, *supra* note 34, R. 1.5). Second, "[i]n addition to the rules which by their terms leave room for moral deliberation, the drafters of the disciplinary rules correctly note that in addition to substantive and procedural law, 'a lawyer is also guided by personal conscience

and the approbation of professional peers.'" *Id.* (quoting MODEL RULES, *supra* note 34, pmbl.). As an example of the second category, Wendel cites the approach of the Restatement, adopted in the amended Model Rule 1.6 as well, permitting but not requiring disclosure of confidential information to prevent death or serious bodily harm. *See id.* at 11-12; *supra* Part II.C.

Although the present analysis centers on the relevance of the Deliberative Model to the second category, and indeed the same example, that Wendel identifies, the Deliberative Model is relevant to the interpretation and application of rules included in the first category as well. In effect, the Deliberative Model looks to extend to discretionary rules that fit within the second category an interpretive methodology that has already been applied to the rules that fall under the first category.

Professor Zacharias has likewise identified a number of ethics rules that "afford lawyers significant discretion in accepting clients, giving advice, and selecting litigation tactics." Zacharias, *supra* note 1, at 1326. For example, Zacharias finds ethics codes "ambiguous" in relation to tactics employed in the course of negotiations: The Model Rules "forbid making false statements in negotiations, while condoning 'puffing' and adherence to negotiation 'conventions.'" *Id.* at 1335 (quoting MODEL RULES, *supra*, R. 4.1, cmt. [2]). In addition, "[t]he Rules encourage lawyers to communicate with clients regarding appropriate means and objectives, while emphasizing the lawyer's duty to maximize client interests and the client's right to control settlements and pleas." *Id.* at 1135-36 (citing MODEL RULES, *supra* note 34, R. 1.4(b), cmt. 1; R. 1.2(a)). Finally, "[t]he codes express disapproval of delay and state that lawyers need not employ all possible tactics, but also appear to subordinate these preferences to client interests." *Id.* at 1136 (citing MODEL RULES, *supra* note 34, R. 1.3, cmt. 1; R. 1.2, cmt. 1; R. 3.2). On the basis of these examples, Zacharias concludes that "[i]n practice, therefore, the codes provide authority for virtually any approach the negotiating lawyer chooses to take." *Id.*

Indeed, Zacharias makes similar observations in relation to a number of other areas of legal practice, including the "courtroom context," *id.* at 1338-39, the "media context," *id.* at 1339-1340, "employing proper tactics to help wrongful clients win," *id.* at 1340-43, "employing proper tactics to help clients commit new wrongful acts," *id.* at 1344-46, and "employing marginal tactics." *Id.* at 1346-48. *See also id.* at 1329-30 & nn.87-88.

Moreover, Zacharias observes, although—or, more likely, because—"the codes in fact accord lawyers significant choice in selecting tactics, screening arguments, and presenting accurate versions of the facts, . . . lawyers have come to use the model of client-oriented ethics as a shield, both for defending behavior and for avoiding introspection regarding moral issues." *Id.* at 1349. *See also* Zacharias, *Specificity*, *supra* note 51, at 263 (asserting that "the requirement that 'prosecutors should do justice' . . . enables prosecutors to justify virtually any response to any ethical dilemma" and that as a result, nearly "anything they decide is professionally correct"); Zacharias, *supra* note 140, at 1942-43 (asserting that "[l]awyers have more influence on the conduct of lawsuits than they are ready to admit," and concluding that "[f]or as many types of conduct as the codes require of lawyers, there are more instances in which the codes impose permissive duties to serve societal interests or in which the codes grant lawyers ultimate decisionmaking discretion").

Cf. Hodes, *supra* note 62, at 1558 (noting that "discretionary language still abounds in the Model Rules, recognizing—while at the same time underscoring—the continuing need for judgment"); Stier, *supra* note 20, at 584 (stating that "much of both the [Model] Code and the [Model] Rules provide lawyers with opportunities to decide for themselves whether or not to take a particular action"); *id.* at 596 & n.179 (noting that "[m]uch of the law of lawyering makes room for morals by giving lawyers discretion in determining what they ought to do," not only when "[t]he

The significance of the Deliberative Model may be considered in light of its application to the rules of confidentiality. Although it shares with the models proposed by Simon and Strassberg the goal of increased and improved ethical deliberation, the Deliberative Model differs significantly from both of these models, as demonstrated in a comparison of the different approaches to the scenarios of the Innocent Convict and *Spaulding*.

In applying the Deliberative Model to these scenarios, the analysis first acknowledges that in both cases the applicable rule—Model Rule 1.6—states that

---

language of some standards is deliberately couched in permissive rather than mandatory terms," but also when "[e]ven mandatory standards . . . leave room for judgment calls").

The Deliberative Model is intended, in part, as an effort to respond to and remedy this phenomenon, requiring that lawyers engage in ethical introspection to arrive at an ethically justifiable decision under a particular set of circumstances, even when involved in an area of practice in which ethics codes do, in fact, afford discretion to lawyers rather than mandating specific conduct. It may not be surprising, in light of his analysis, that one of the remedies Professor Zacharias proposes shares a number of similarities with—but nevertheless differs significantly from—the Deliberative Model. *See* discussion *infra* note 154.

Moreover, some scholars have suggested not only that, as a descriptive matter, "the current Model Rules and Model Code do not even aim to cover all ethically required conduct, let alone all ethically exemplary conduct," Feldman, *supra* note 1, at 902 n.55, but that, as a jurisprudential matter, there exists an "inevitable inconclusiveness inherent in any statutory code, including a code of ethics." *Id.* at 898. Professor David Wilkins' groundbreaking work in this area applies to legal ethics each of three primary arguments set forth by "legal realists and their followers" to support the claim that "law is largely indeterminate." Wilkins, *supra* note 44, at 478.

Specifically, Wilkins considers the relevance to the practicing lawyer of the arguments that: (1) "there are in most cases a number of sources from which a 'legal' answer might be derived"; (2) "legal doctrines contain vague or ambiguous language susceptible to multiple, inconsistent interpretations"; and (3) "by shifting the focus of analysis between the general and the particular, it is often possible to alter the perception of the proper application of the law to the facts." *Id.* at 478-84. In support of the second of these propositions, Wilkins observes that the Model Code is "rife with vague and ambiguous terms," *id.* at 480 (citing MODEL CODE, *supra* note 34, DR 1-105(A); DR 7-102(A)(2); DR 9-101), and that, although the tone and structure of the Model Rules "eliminated some of the more pervasive ambiguities, vagueness and open-endedness remain." *Id.* at 480-81 (citing MODEL RULES, *supra* note 34, R. 1.1; R. 3.4(d)).

Finally, both courts and legal scholars have emphasized that, "[a]s a practical matter, there could never be a set of rules which contemplates every aspect of the many encounters between an attorney and client." *In re* Rinella, 677 N.E.2d 909, 917 (Ill. 1997). *See* Levine, *supra* note 36 (citing and analyzing sources). Indeed, as I have noted elsewhere, courts and scholars have recognized a similar impracticality in articulating every right guaranteed by the United States Constitution or, under Jewish legal theory, every obligation mandated by the Torah. *See id.* I have suggested that, on both a descriptive and normative level, all three of these areas of law adopt a hermeneutic methodology through which broad provisions are interpreted as a basis for deriving and identifying otherwise unenumerated ethical obligations, constitutional rights, and biblical obligations, respectively. *See id.*

a lawyer "may" reveal confidential information necessary to save a life.[142] Yet, because under the Deliberative Model this Rule is seen as discretionary rather than optional, the analysis does not end with the simplistic and categorical conclusion that, regardless of circumstances and even in the absence of any ethical deliberation, a lawyer may ethically reveal—or not reveal—the information. Instead, according to this analysis, in most—or, quite likely, as the rule implies, all—cases in which the lawyer faces such an ethical dilemma, the lawyer *may* be ethically permitted either to disclose or not disclose the information. Nevertheless, upon review of the action taken, the lawyer may be required to justify the ethical quality of the decision, through an articulable demonstration of ethical deliberation supporting the decision.[143]

The difference between the Deliberative Model and Simon's model is apparent in the context of the case of the Innocent Convict. According to Simon, even under the former Rule 1.6, which, by virtually all accounts, would have flatly prohibited disclosure[144]—and, *a fortiori*, under the current, more permissive rule—disclosure is considered not merely an optional decision for the ethical lawyer but the one correct decision.[145] In contrast, under the Deliberative Model, although specific circumstances and various forms of deliberation may likewise support the ethical decision of disclosure, the individual lawyer retains

---

142. *See* MODEL RULES, *supra* note 34, R. 1.6(b)(1).

143. The Deliberative Model thus recognizes that different—possibly contradictory—responses may be ethically proper in the same situation, as long as each is justifiable on the basis of ethical deliberation. *See* Atkinson, *supra* note 75, at 1515 (advocating "healthy toleration of ethical diversity" based on the principle that "[b]eyond certain widely agreed minimal standards, conscientious lawyers genuinely and perhaps irreducibly disagree about what their ethical obligations are," insisting that "[i]n the face of that kind of disagreement, a well-meaning majority might well opt to allow legal latitude to the loyal opposition, even as the majority holds out its view in an official form as ethically superior," and invoking a "religious parallel" to note that "one can have an established church without persecuting dissenters"); RHODE, *supra* note 43, at 58 ("Lawyers can, and should, act on the basis of their own principled convictions, even when they recognize that others could in good faith hold different views"); Zacharias, *Specificity*, *supra* note 51, at 258 (arguing that "in guiding lawyers rather than directing . . . particular acts, the codes acknowledge that there may be more than one appropriate response to the situations in question."); *cf.* W. Bradley Wendel, *Value Pluralism in Legal Ethics*, 78 WASH. U. L.Q. 113, 116-17 (2000) (observing that "the foundational normative values of lawyering are substantively plural and, in many cases, incommensurable" and prescribing that, therefore, "the model of ethical decision making for lawyers accommodate the incommensurability of professional values and make appropriate adjustments.").

Thus, under the Deliberative Model, as under Simon's model, the decisionmaking process of the private lawyer approximates that of the prosecutor, whose obligation to do justice "may . . . point in contradictory directions." Green, *supra* note 49, at 622. *See* SIMON, *supra* note 1, at 10-11; *see also* Zacharias, *Specificity*, *supra* note 51, at 250 ("Different prosecutors . . . can justify diametrically opposite conduct as serving justice.").

144. *See supra* note 115.

145. *See supra* Part II.A.

the discretion not to disclose when such a decision is based in demonstrable ethical deliberation.[146]

At the same time, if challenged, the lawyer who chooses disclosure to save the life of the Innocent Convict may be required to demonstrate ethical considerations supporting such a decision under the specific circumstances of the case.[147] Thus, when compared with Simon's model, which prescribes disclosure

---

146. The Deliberative Model likewise differs from the approach prescribed by California rules of confidentiality, flatly prohibiting disclosure of confidences to save a life. As Professor Zacharias has developed at some length, a reasonable interpretation of California's generally categorical laws against disclosing confidences may have permitted disclosure in the extreme case of a client's stated intent to commit a murder. *See* Fred C. Zacharias, *Privilege and Confidentiality in California*, 28 U.C. DAVIS L. REV. 365 (1995). Indeed even the generally restrictive Model Code and former Model Rule 1.6 include provisions permitting disclosure under such a scenario. *See* MODEL CODE, *supra* note 34, DR 4-101(C)(3); GILLERS & SIMON, *supra* note 63, at 70. According to the interpretation set forth in a California ethics opinion, however, the lawyer must nevertheless maintain the client's confidences. *See* Zacharias, *supra* note 1, at 1329 n.86 (citing San Diego County Bar Ass'n Legal Ethics and Unlawful Practices Comm., Op. 1990-1) (applying CAL. BUS. & PRO. CODE § 6068(e)).

Although the outcome of this interpretation of the California rule is directly contrary to Simon's prescription of disclosure to save a life, the bar committee similarly shares with Simon a methodological framework that mandates a single ethically acceptable response for lawyers. *See id.* (concluding that, under such an interpretation, the California rule "depriv[es] lawyers the right or obligation to weigh values" by "establish[ing] an absolute principle that a lawyer's role in the system includes a duty not to weigh considerations militating against confidentiality"). In contrast to both of these approaches, the Deliberative Model considers either disclosure or silence a potentially proper ethical response, mandating instead that the lawyer engage in ethical deliberation in deciding how to exercise discretion. *Cf.* Reed Elizabeth Loder, *Tighter Rules of Professional Conduct: Saltwater for Thirst?*, 1 GEO. J. LEGAL ETHICS 311, 316 (1987) ("Prepackaged solutions remove situational ambiguities and the need to make and justify difficult choices, two features which characterize moral decisionmaking.").

147. The approach of the Deliberative Model in the case of the Innocent Convict may also be elucidated through a contrast to the Restatement. Like the current Model Rule 1.6, the Restatement permits disclosure of confidential information to prevent death or serious bodily injury. RESTATEMENT, *supra* note 36, § 66(1). Unlike the Deliberative Model, however, the Restatement considers the decision to disclose or not disclose to be optional rather than discretionary. Accordingly, under the Restatement, "[a] lawyer who takes action or decides not to take action permitted under this Section is not, solely by reason of such action or inaction, subject to professional discipline." *Id.* § 66 (3).

The Comment to this Section explains the rationale behind the optional approach to disclosure. Specifically, the Comment observes that the decision to disclose "would inevitably conflict to a significant degree with the lawyer's customary role of protecting client interests." *Id.* at cmt. g. Moreover, "[c]ritical facts may be unclear, emotions may be high, and little time may be available in which the lawyer must decide on an appropriate course of action." *Id.* Therefore, the Comment continues, "[s]ubsequent re-examination of the reasonableness of a lawyer's action in light of later developments would be unwarranted." *Id.* Instead, "reasonableness of the lawyer's belief at the

as a mandatory ethical decision, the Deliberative Model would appear both to increase and to require the exercise of ethical discretion, through ethical deliberation on the part of individual lawyers.[148]

---

time and in the circumstances in which the lawyer acts is alone controlling." *Id.*

A close look at the reasoning of the Comment suggests a logical flaw, or at least a leap in the logic—in effect, a non-sequitur—from these convincing premises to the Restatement's more questionable conclusion that the lawyer's decision to disclose or not disclose is not subject to disciplinary review. The Comment first asserts the premise—undoubtedly correct—that the decision of an attorney to disclose a client's confidential information, even to save a life, is a difficult one, both because it entails action contrary to the goals of the client and because it involves complex ethical judgment under circumstances that may not be conducive to careful ethical deliberation. Based on the initial premise, the Comment further posits—again, rather convincingly—that any consideration of the reasonableness of the lawyer's decision must be evaluated exclusively upon the lawyer's belief at the time of and under the circumstances of the lawyer's actions, rather than in light of later circumstances.

At this point, the Restatement's logic would not seem substantially different from the approach of the Deliberative Model. Indeed, the Deliberative Model emphasizes the complex nature of the decision whether to disclose confidences, acknowledging that differing circumstances and considerations may suggest different responses. As a result, the Deliberative Model incorporates both substantive and procedural limitations on the extent to which a lawyer's exercise of discretion to disclose or not disclose may later be questioned. *See* discussion *infra*. In fact, the approach of the Deliberative Model is thus consistent with the statement in the Comment that "[s]ubsequent re-examination . . . of the lawyer's action in light of later developments would be unwarranted." RESTATEMENT, *supra* note 36, cmt. g.

Unlike the Deliberative Model, however, the Restatement goes a step further, reaching a different conclusion, without apparent support in the logic of the premises on the basis of which this conclusion is ostensibly constructed. The Restatement goes beyond the concern that comprises the focus of the comment—the need to protect the lawyer's decision from second-guessing based on incomplete and inappropriate factors. Instead, the Restatement precludes any form of review, thereby essentially adopting an optional rule and obviating any requirement of ethical deliberation.

Interestingly, the introductory Scope of the Model Rules formerly included the statement: "The lawyer's exercise of discretion not to disclose information under Rule 1.6 should not be subject to reexamination." *See* GILLERS & SIMON, *supra* note 63, at 13 (quoting MODEL RULES, *supra* note 34, Scope [20] (now deleted)). It is not clear why this provision was initially adopted as part of the Scope of the Model Rules, as it appears merely to repeat one aspect of the general principle that a lawyer's decision to act or not act pursuant to a permissive rule is not subject to disciplinary action. Conversely, there does not appear to be any significance to the deletion of this provision, as its deletion likewise does not affect the same general principle precluding discipline for decisions undertaken in relation to permissive rules. Perhaps the inclusion—and later deletion—of this provision indicates at least a recognition, even within the Model Rules, that on some level discretionary decisions regarding confidentiality are different from other permissive rules.

148. In addition, unlike Simon's solution, the Deliberative Model thus acknowledges that there may be more than one ethically justifiable response to the case of the Innocent Convict, as reflected in the policy decision of the ABA to amend the Rule to permit—but not to categorically mandate—disclosure. *See supra* note 118; Cramton & Knowles, *supra* note 64:

Concomitantly, the Deliberative Model differs from Strassberg's model, as illustrated in an analysis of the scenario in *Spaulding*. When applied to *Spaulding*, Strassberg's model would certainly encourage disclosure, but apparently would require neither a specific outcome nor a justification for a decision either to disclose or not disclose.[149] According to the Deliberative Model, while either of these decisions may likewise be available to the individual lawyer, it would be insufficient—if not inaccurate—to state merely that the lawyer has the "option" to disclose or to maintain the client's confidence.

Instead, under the Deliberative Model, the lawyer has the discretion to make an ethical determination, accompanied by an obligation to engage in ethical deliberation before arriving at one of the possible conclusions. Therefore, unlike Strassberg's model, which may be understood to involve ethical deliberation of the lawyer only to the extent that the individual lawyer voluntarily undertakes such deliberation, the Deliberative Model requires the individual lawyer to exercise ethical discretion through ethical deliberation.[150]

> The variety and uniqueness of the circumstances that must be considered confirm our preference that, as a general matter, exceptions to confidentiality be cast in discretionary terms. Broad legal commands are unlikely to reflect the moral complexity of many real-life situations. The lawyer must consider the unique characteristics of the individual case as well as its consonance with values held dear by the community.

Id. at 119; Hodes, *supra* note 62, at 1561 (noting that "the specific context of each case—including the [Innocent Convict]—colors what can be done to alleviate the tragedy, which affects what ought to be done" and concluding "[t]hat focus should, in turn, significantly affect the judgments we make in assessing what the lawyers actually did do in any particular case."). *See also* Feldman, *supra* note 1, at 932-33 (stating that "one of the reasons for the genuine difficulty of live ethical problems is the uniqueness of such cases" and that "[t]he distinctiveness of the configuration stems from the specifics of the circumstances"); Strassberg, *supra* note 1, at 951-52 n.257 (suggesting an approach "compatible with the numerous attempts to broaden permissive disclosure by attorneys which would allow individual attorneys to weigh the principles supporting and opposing disclosure in particular fact situations in which a great deal of controversy exists over what is ethical conduct").

149. *See supra* Part II.B.

150. The Deliberative Model thus requires lawyers to exercise judgment in a manner that Professor Stier attributes to the "good reader" of discretionary rules, rather than adopting the decision-making process of the "bad reader." *See* Stier, *supra* note 20, at 594. Stier considers these two categories of lawyers in the context of a rule permitting the disclosure of confidences to prevent a client's future crime that will result in substantial bodily harm to another. In Stier's framework, bad readers "determine[] whether it is prudent to do nothing that is not required or whether their self-interest lies in some actions and what that may be." *Id.* Accordingly, if the bad reader decides to remain silent, the "reasons for doing so would be bad reasons like concern for losing clients and fees." *Id.*

Good readers, in contrast, "are not foreclosed from acting prudently, but the judgments they must make require even more. They must also integrate and balance possibly competing legal, moral, and prudential reasons for action." *Id.* For example,

> [t]he good reader who defends drug sellers might justify . . . silence by reference to some moral value like the centrality of confidentiality to maintaining a relationship with

Of course, like virtually any proposed model of reform, the Deliberative Model may encounter criticisms and contain weaknesses. Perhaps the most fundamental objection that may be lodged against the Deliberative Model is the concern that, because it requires a lawyer to justify an ethical decision even in the case of a discretionary rule—and, in the absence of such articulable justification, may subject the lawyer to discipline—it curtails the range of discretion available to a lawyer. A defense of the Deliberative Model against such an objection may be offered on at least two different levels, one substantive and the other procedural.

On a substantive level, admittedly, requiring justification of ethical decisions may, in some cases, limit the range of actions available to the lawyer. It is indeed possible that a lawyer may be deterred from making a decision that could lead to acting in a manner that may be ethically proper—or perhaps ethically preferable—because of the lawyer's inability or unwillingness to engage in and articulate ethical deliberation in support of such a decision. However, as scholars have observed, as a direct result of the current approach, which views discretionary rules as optional rules, "lawyers have come to use the model of client-oriented ethics as a shield, both for defending behavior and for avoiding introspection regarding moral issues."[151] The Deliberative Model directly

---

> the criminal client and affording that client a fair representation under the conditions provided by the current adversary system of justice.

Id.

Significantly, like the Deliberative Model, Stier's analysis recognizes that two lawyers may decide upon the same course of action through two very different processes of decisionmaking. Moreover, in such a case, the ethical quality of their judgments is dependent on a complex consideration of the nature of their ethical deliberations, not on a superficial consideration of the outcome of the decisionmaking process.

A similar emphasis on the quality of the process of ethical deliberation is central to the question Professor Feldman's asks in the title of her article, *Can Good Lawyers be Good Ethical Deliberators?* Feldman, *supra* note 1. *See id.* at 929 ("Primarily, I am interested in the character of the lawyers' ethical deliberation, rather than whether their actions were ethically correct."); *id.* ("I do not mainly measure the authenticity of the lawyers' ethical deliberations according to the choices they ended up making."). Likewise, a student who participated in a seminar of the same name offered by Feldman concluded that "it is not only the outcome, but also the reasoning process itself, that constitutes good ethical deliberation." Stephanie Loomis-Price, Note, *Decision-Making in the Law: What Constitutes a Good Decision—the Outcome or the Reasoning Behind It?*, 12 GEO. J. LEGAL ETHICS 623, 624 (1999).

151. Zacharias, *supra* note 1, at 1349. *See also* Feldman, *supra* note 1, at 898 ("The more frequently a black letter ethics code is inconclusive, the more opportunities there are for . . . interpreting the rules simply to permit pursuit of the client's ends, without regard to independent ethical concerns.").

Moreover, as Professor Wilkins has noted, "the traditional model strongly implies that doubts about the exact contours of the law should be resolved in the client's favor." Wilkins, *supra* note 44, at 473. *See id.* at 473 n.17 (citing MODEL CODE, *supra* note 34, DR 7-101 (A) (1); EC 7-4; EC 7-5; MODEL RULES, *supra* note 34, R. 3.1, cmt.[1]) (adding that "[t]he rules of professional conduct

confronts each of these problems, requiring ethical deliberation and thereby preventing lawyers from blithely justifying nearly any conduct through assertions of acting in the client's interests.[152] Thus, despite occasional—if not

---

generally support the view that all doubts should be resolved in favor of furthering the best interest of the client."). The ramifications of this approach to decisionmaking in the face of doubts—including the problems that seem to inhere in such an approach to legal ethics—may be examined through another comparison to Jewish law.

The complexities of resolving legal questions when there remains an insoluble doubt find extensive expression in the corpus of Jewish law. *See, e.g.*, ARYEH LEB HA-COHEN HELLER, SHEV SHEMAT'TA. As a general rule, an even doubt relating to the applicability of an obligation or commandment of biblical authority is resolved in favor of requiring conscious adherence to the obligation. *See id.* This principle would appear broadly parallel to the notion of resolving ethical doubts in favor of adherence to the interests of the client.

Nevertheless, Jewish law recognizes that, at times, a counterbalancing interest may override adherence even to obligations of biblical authority. Specifically, nearly every obligation in Jewish law is suspended to save a life. *See* MAIMONIDES, MISHNE TORAH, LAWS OF SABBATH, ch. 2 (Rabbi Eliyahu Touger trans.) (1993). In fact, the primacy of saving a life is such that "[e]ven if there is only a doubtful possibility that a person's life is in danger, one renders a lenient decision [to violate other obligations]; and as long as one is able to discover some possible danger to life, one may use that doubt to render a lenient decision." JOSEPH B. SOLOVEITCHIK, HALAKHIC 34-35 (Lawrence Kaplan trans., 1983) (originally published in Hebrew as *Ish ha-halakhah*, in 1 TALPIOT 3-4 (1944)). Thus, while a minor possibility of the applicability of an obligation generally does not mandate adherence to the obligation, even a minor possibility of danger to life justifies—and requires—any necessary response, including violation of nearly any obligation. *See id.* at 34 (quoting MAIMONIDES, *supra*, *Laws of Sabbath* (2:3)) ("[I]t is forbidden to delay such violation of the Sabbath for the sake of a person who is dangerously ill.").

In contrast, the mechanical approach to legal ethics documented by Wilkins fails to acknowledge the complexity of ethical decisionmaking, relying instead on adherence to client goals as an avenue for avoiding meaningful consideration and balancing of competing interests.

152. In fact, it is arguable that, rather than restricting the range of conduct available to lawyers, the requirement that lawyers engage in ethical deliberation may expand the range of possible action, compelling lawyers to consider responses they otherwise might have reflexively rejected as inconsistent with their own personal/economic interests and/or those of their clients. Indeed, under the current model of ethics regulation,

> [w]hen the codes authorize lawyers to choose between emphasizing partisanship and important third party or societal interests, lawyers' natural [personal and economic] incentives encourage them to select partisanship. Lawyers who make that choice can readily justify their conduct as mandated by the code by claiming adherence to the code provisions that call for zeal.

Zacharias, *supra* note 1, at 1340. *See also id.* at 1304 (observing that "[l]awyers have obvious economic incentives to pursue client desires aggressively" and that "[l]awyers typically also feel comfortable allying themselves with clients whom they know personally and who may exert an influence on their well-being.").

Professor Zacharias has collected a number of common statements lawyers offer to justify "hardball tactics" and "decisions to assert questionable claims or arguments or to delay litigation,"

rare—practical drawbacks, the Deliberative Model offers a preferable framework through which lawyers are generally encouraged to consider viable ethical alternatives as they are required to articulate justifications for their ethical decisionmaking.[153]

---

or in response to "the claim that the lawyer has misled a jury or court." *Id.* at 1348-49. As Zacharias observes, "[e]ach of these statements suggests that," as a result of a duty of zealous loyalty to clients' interests, "the lawyer had no discretion to act differently. . . . [t]hat the ethics of the profession—the codes—-required [the] conduct." *Id.* at 1349. If, as this analysis indicates, lawyers feel—or at least contend—that under many circumstances they are prevented from exercising free choice of ethical conduct, while perhaps counterintuitive, it may be necessary to impose an obligation of ethical deliberation in order to return to lawyers such freedom of ethical decisionmaking.

To draw another parallel to Jewish law and philosophy, a number of leading Medieval scholars grappled with the question of the apparent contradiction between the fundamental principle in Jewish thought that human beings exercise free will and the biblical verses stating that, at certain points, God "hardened Pharaoh's heart" to prevent Pharaoh from deciding to free the Nation of Israel from slavery in Egypt. *See, e.g.*, *Exodus* 7:13; 7:22; 8:15; 9:12; 9:35; 10:20; 10:27; 14:7; 14:8. Of course, the issue of human free will is extraordinarily complex and has captured the attention of philosophers since antiquity. *See, e.g.*, Sherman J. Clark, *The Courage of Our Convictions*, 97 MICH. L. REV. 2381, 2402 n.39 (1999) (citing sources); Samuel J. Levine, *Playing God: An Essay on Law, Philosophy, and American Capital Punishment*, 31 NEW MEX. L. REV. 277, 286-89 (2001) (discussing the issue of free will in Jewish law and philosophy).

Regarding the question of Pharaoh's free will, Rabbi Yoseph Albo offers a response that may help illuminate one of the functions of the Deliberative Model in improving the quality and range of ethical decisionmaking. Rabbi Albo explains that, as a result of the miraculous plagues that God brought against Egypt, Pharaoh's free will had effectively been removed; in the face of such Divine intervention, it would have been impossible for a human being freely to choose any response other than to send the nation out of bondage. *See* YOSEPH ALBO, SEFER HA'IKKARIM (4:25). Therefore, such a decision by Pharaoh would not have been the product of free will. Accordingly, God influenced Pharaoh's decisionmaking process to the extent of removing this impediment to Pharaoh's exercise of free will, thereby in fact returning to Pharaoh the discretion to decide whether to free the nation. *Id.* Similarly, by requiring ethical deliberation, the Deliberative Model aims in part to remove impediments to careful ethical decisionmaking often posed by the natural tendency of lawyers to avoid the exercise of discretion in favor of simplistic adherence to personal and economic pressures.

For other responses to the specific question of Pharaoh's free will, see, e.g., MAIMONIDES, *supra* note 151, *Laws of Teshuva* (6:3); MAIMONIDES, INTRODUCTION TO COMMENTARY ON THE MISHNA, *Introduction to Pirke Avoth*, ch. 8; 1 NACHMANIDES, COMMENTARY ON THE TORAH 309 (Chaim Chavel ed., 1960) (commenting on *Exodus* 7:3).

153. Ethics scholars have observed that "institutional justifications for actions are 'seductive' because they limit the lawyer's decisionmaking within a structured and 'simplified moral world,'" Loder, *supra* note 146, at 315-16 (quoting Richard Wasserstrom, *Lawyers as Professionals: Some Moral Issues*, 5 HUM. RTS. 1, 8 (1975)), and that, therefore, "the rules may provide an escape from an otherwise painful process of justifying moral choices." *Id.* at 319.

This psychological insight evokes yet another parallel to Jewish thought. The biblical

Moreover, on a procedural level, the Deliberative Model may avoid improper curtailment of ethical discretion through the adoption of a standard of review for discipline that takes into account the range of ethically justifiable conduct that may apply under a given set of circumstances: the abuse of discretion standard. The utility of this standard for reviewing ethical decisions of lawyers goes beyond semantic similarities. The abuse of discretion standard is an important device through which appellate courts have the ability, in reviewing certain types of rulings, to maintain appropriate supervision of trial court decisions while according trial courts a substantial and deserved measure of discretion.[154]

---

narrative of the binding of Isaac, *see Genesis* 22:1-19, begins with the statement that "God tested Abraham." *Id.* at 22:1. Ostensibly, the nature of the test was whether Abraham would abide by the apparent command to sacrifice "your son, your only one, whom you love," *id.* at 22:2, and whose birth Abraham had awaited for so many years. *See id.* at 17:15-21; 21:1-7.

Alternatively, however, some offer an innovative and psychologically complex interpretation of the test facing Abraham. Specifically, a close look at the biblical text suggests that Abraham understood God's command to sacrifice Isaac only after close and careful interpretation of God's arguably ambiguous words. *See* RABBI MORDECHAI YOSEPH OF IZHBITZ, MEI HA-SHILOACH (explicating *Genesis* 22:1). Thus understood, the test facing Abraham was the obligation to engage in honest and painful moral deliberation resulting in the decision to undertake the unimaginable action of sacrificing his son. *See id.*; Rabbi Menachem Mendel Blachman, *The Sacrifice of Intellectual Honesty*, *available at* http:// www.kby.org. On some level, this was a more painful decision than adhering to an unambiguous command, however difficult and demanding. *See also* RABBI YOSEPH YOZEL HURWITZ, MADREGAT HA-ADAM 226-31 (4th ed. 1976) (describing human tendency to prefer simplistic and self-interested solutions in place of complexity); RABBI CHAIM SHMULEVITZ, SICHOTH MUSAR 95 (1980) (describing human tendency to prefer definite result or command—even negative or burdensome one—rather than remaining in state of uncertainty).

154. The Supreme Court has described deference to the trial court as "the hallmark of abuse-of-discretion review," emphasizing that "such deference [is] owed to the judicial actor . . . better positioned than another to decide the issue in question" and that "deferential review . . . afford[s] the district court the necessary flexibility to resolve questions involving multifarious, fleeting, special, narrow facts that utterly resist generalization." *See, e.g.*, Koon v. United States, 518 U.S. 81, 98-99 (1996). The Court emphasized that such "deference [is] owed to the judicial actor . . . better positioned than another to decide the issue in question" and that "deferential review . . . afford[s] the district court the necessary flexibility to resolve questions involving multifarious, fleeting, special, narrow facts that utterly resist generalization." *Id.* at 99 (internal citations and quotations omitted). Likewise, the Deliberative Model acknowledges the deference owed to the lawyer who exercises professional judgment and who should be afforded flexibility in resolving complex, fact-specific ethical dilemmas.

Indeed, an important element of the Deliberative Model is its ability to withstand objections that place limitations on other models of ethics regulation, through the incorporation of both substantive and procedural principles applied in the American legal system in contexts other than legal ethics. For example, the Deliberative Model shares a number of similarities with a proposal by Professor Zacharias "[c]odifying the [d]uties [t]o [e]ngage in [m]oral [d]iscourse and [i]ntrospection." Zacharias, *supra* note 1, at 1357. A fundamental premise of the Deliberative Model, echoing the approach offered by Zacharias, is its emphasis that it "need not prescribe

particular outcomes" but that nevertheless it "can encourage lawyers to take the duty of objectivity more seriously." *Id.* at 1360-61. Moreover, under both models, "the lawyer would need to document that she has considered the moral issues. . . ." *Id.* at 1367-68.

Nevertheless, Zacharias restricts his proposal to an obligation on the lawyer to "discuss[] the limits on advocacy with her client . . . and discuss[] the particular deposition and deposition strategy with the client, including the appropriateness of harassing or intimidating the plaintiff." *Id.* at 1368. The proposal thus places a significant limitation on the duty of ethical discourse and introspection, restricting the requirement to conversations with clients. According to Zacharias, restricting the review by courts to what he considers "verifiable action by lawyers," *id.* at 1367, is necessary for two apparently interrelated yet separate reasons, both based in the recognition that "[a] requirement of introspection, by definition, is difficult to enforce." *Id.* First, "[d]isciplinary authorities cannot know what lawyers 'have thought.'" *Id.* Second, "[u]pon questioning, lawyers can rationalize most conduct after the fact." *Id. Cf.* Stephen Gillers, *More About Us: Another Take on the Abusive Use of Legal Ethics Rules*, 11 GEO. J. LEGAL ETHICS 843, 846 (1998) (positing the "near-impossibility of proving the lawyer's 'true' motive").

Under the Deliberative Model, however, the obligation of introspection—or ethical deliberation—applies to the exercise of any discretionary decision by the lawyer, including when such deliberation is not undertaken through communication with another person. As for the concerns about verifiability voiced by Zacharias, the Deliberative Model can again rely on procedures common to the American legal system and their underlying assumptions.

While it may be impossible to know with certainty the thoughts of an individual, our legal system—in particular the criminal law—is replete with rules that depend on the motive or state of mind of an individual. The law accepts—as it must—circumstantial evidence to establish the state of mind for intentional crimes, justifying punishment on the basis of such evidence. *See, e.g.*, United States v. Nelson, 277 F.3d 164, 197 (2d Cir. 2002) ("[I]t is well-settled that, as a general matter, criminal intent may be proven by circumstantial evidence.").

Likewise, as Zacharias himself emphasized in an earlier article, in a number of contexts, "[c]ourts have coped with problems of determining the motives of government administrators" and "[i]t is [thus] the task of judges to infer intent from extrinsic circumstances, corroborating documentary evidence, and the demeanor of witnesses." Fred C. Zacharias, *Flowcharting the First Amendment*, 72 CORNELL L. REV. 936, 977 & nn. 207-08 (1987). *See also* Michael Selmi, *Proving Intentional Discrimination: The Reality of Supreme Court Rhetoric*, 86 GEO. L.J. 279, 290 (1997) ("emphasiz[ing] that the need to rely on circumstantial evidence to prove [discriminatory] intent is not a new issue"). Therefore, it would not seem unreasonable to rely on circumstantial evidence, when necessary, to review the ethical deliberations and decisions of lawyers.

Moreover, the legal system—again, in large part out of necessity, it might seem—allows individuals to attempt to rationalize conduct after the fact. For example, in preparing for a criminal trial, a prosecutor may have the task of deriving a theory justifying the tactics the police employed in apprehending a defendant, while conversely, a defense attorney may be faced with the prospect of formulating a theory justifying the defendant's apparently criminal conduct. The fact that attorneys may find a way to justify their own conduct when subject to review should be no more problematic than their ability to perform a similar service for those they represent and defend.

In establishing the Contextual View, Simon similarly aimed to offer a model that would be consistent with general principles of the legal system, emphasizing its reliance on the principles of justice and contextual judgment and the prevalence of these principles in modern jurisprudence.

Similarly, application of the abuse of discretion standard to the ethical decisionmaking of lawyers strikes a proper balance, requiring lawyers to engage in articulable ethical deliberation[155] without improperly curtailing the range of

---

In complying with principles and procedures already established in the legal system, the Deliberative Model may succeed in meeting such a goal.

More generally, the Deliberative Model may satisfy one of the criteria some scholars have set for alternative models, as it may be readily applied under the current structure of ethics regulation. *See, e.g.*, Strassberg, *supra* note 1, at 904, 923 (concluding that "[i]f we are to have legal ethics, it must be compatible with the existence of ethical rules" and accordingly proposing a model of "interpretation and application of the existing rules of ethics"); Zacharias, *supra* note 140, at 1942 (asserting that "it is not necessary to trash existing professional standards to sanction tamer lawyering 'in the public interest'" and that "[t]he current codes already allow it"); *cf.* Shannan E. Higgins, Note, *Ethical Rules of Lawyering: An Analysis of Role-Based Reasoning from Zealous Advocacy to Purposivism*, 12 GEO. J. LEGAL ETHICS 639, 662 & n.67 (1999) (suggesting an approach that does not "suggest a complete overhaul of the legal profession," acknowledging that the approach involves a "change" that may not "be easy," but concluding that "manipulation of the present legal and ethical discourse might be a more productive course of action than . . . styling a new theory of legal ethics"). Although it suggests a new interpretive framework, the Deliberative Model is generally consistent with the substance of existing rules, their underlying goals, and the mechanism employed for their enforcement.

155. Thus, by incorporating a degree of objectivity, this standard may be consistent with Professor David Luban's insistence that "legal ethics isn't only a matter of a lawyer's first-personal belief: The belief needs some plausible basis." Luban, *supra* note 53, at 891. *Cf.* Loder, *supra* note 146, at 330 (noting that "permissive rules may perpetuate a public perception that lawyers are oftentimes free from institutional accountability"); Wilkins, *supra* note 44, at 524 ("Legal ethics owes the profession and society a credible account of how [lawyers'] discretion should be exercised.").

Professor Zacharias has articulated and applied a similar observation in an analysis of the form of ethics rules, positing that although "[o]rdinarily, emphasizing discretion implies that two or more responses to a situation are equally good[,] . . . [t]here may be correct or incorrect responses to a particular situation, given the priorities set in the rules." Zacharias, *Specificity*, *supra* note 51, at 246 n.74. Nevertheless, he continues, drafters "may rely on a [] provision [that suggests but does not require a result] because of their inability to predict the situation or to write a useful rule that is sufficiently broad to encompass that *and other* situations." *Id.* (emphasis added). Through its insistence on an articulably justifiable basis for ethical decisionmaking, the Deliberative Model may provide a means of evaluating responses—and identifying "incorrect" responses—even in the context of a discretionary rule that does not require a particular result.

Likewise, the Deliberative Model might have offered a compromise position for a controversial proposed draft of the Restatement of the Law Governing Lawyers that would have immunized a lawyer from malpractice liability for "any action or inaction the lawyer reasonably believed to be required, *or reasonably believed to be discretionary in the judgment of the lawyer, under . . . a professional rule*." *See* Zacharias, *supra* note 1, at 1330 n.88 (quoting Motion Submitted to the ALI (May 1994), amending RESTATEMENT, *supra* note 36, § 76(4) (Tentative Draft No. 7 (per Rule 12.8.5) (1994))) (emphasis added).

Zacharias voiced the concern, apparently shared by the members of the ALI, who rejected the

their discretion by subjecting their decision to possibly inflexible review by others who may prefer an alternative approach.[156] In the absence of a finding of

---

proposed amendment, that such a provision "would enable lawyers to assert virtually any exercise of discretion under the codes as a defense to civil liability." *Id.* In requiring articulable ethical deliberation for actions arising under discretionary rules, the Deliberative Model may provide a response to such concerns, limiting the range of malpractice immunity to discretionary decisions justifiably based in ethical considerations. At the same time, the Deliberative Model would thus offer a framework for promoting the apparent goal of the proposed amendment, extending immunity to discretionary actions that, pursuant to careful and defensible ethical deliberation—although not expressly required—are determined to comprise ethically proper courses of action.

Consistent with its rejection of such an approach, the final version of the Restatement rejected the proposed amendment, instead maintaining the simplistic dichotomy between acts that are "required" and those that are "permissib[le]" in that they are "allow[ed] but . . . not require[d]." *See* RESTATEMENT, *supra* note 36, § 54(1), cmt. h (per Rule 3.5).

156. The Deliberative Model may thereby also satisfy Professor Strassberg's suggestion that the rules "accommodate, within limits, the reasonable judgments of different attorneys." Strassberg, *supra* note 1, at 951-52 n.257. In short, the Deliberative Model recognizes that "[p]rofessional ethics need not look like a branch of science or formal logic in order to provide a satisfying account of how lawyers should resolve practical dilemmas." Wendel, *supra* note 143, at 117.

Indeed, as legal scholars and philosophers have long-emphasized, unlike the exact sciences, legal reasoning is often not reducible to a degree of mathematical precision. Instead, there may often exist more than one viable answer to a legal question. *See, e.g.*, Samuel J. Levine, *Jewish Legal and American Constitutional Theory: Some Comparisons and Contrasts*, 24 HASTINGS CONST. L.Q. 441 (1997):

> [The] Medieval scholar, Nachmanides, . . . noted an inherent difference between legal reasoning and the logic of exact sciences, such as engineering. While it is possible in engineering to prove demonstrably, with mathematical precision, that a particular theory is correct, legal reasoning often involves issues that can be resolved logically in more than one way. The role of a legal interpreter is to examine the evidence motivating each of the possible conclusions, and to determine which conclusion appears most accurate.

Id. at 471 (footnotes omitted). *See also* Owen M. Fiss, *Objectivity and Interpretation*, 34 STAN. L. REV. 739, 744 (1982) ("[T]he meaning of a text does not reside in the text, as an object might reside in physical space or as an element might be said to be present in a chemical compound, ready to be extracted if only one knows the correct process. . . ."); Gerald Graff, *"Keep off the Grass," "Drop Dead," and Other Indeterminacies: A Response to Sanford Levinson*, 60 TEX. L. REV. 405, 410 (1982) ("[T]he practice of interpretation doesn't depend on interpreters' possessing godlike powers to arrive at an 'ultimately provable right answer' that closes the books on further argument about the meaning of a text. Therefore the lack of such godlike power doesn't entail indeterminacy. . . ."); Michael Rosensweig, *Eilu ve-Eilu Divrei Elohim Hayyim: Halakhic Pluralism and Theories of Controversy*, *in* Sokol, *supra* note 137, at 101 (explicating "several talmudic sources suggest[ing] a notion of the inherent value of dissenting views and possibly even of multiple truths").

If a range of plausible interpretations may be available in response to a difficult legal question, it should not be surprising that a range of reasonable approaches may be offered in response to a complex ethical dilemma. *Cf.* Wendel, *supra* note 5:

> One need not endorse a radical critique of the rule of law to express doubts about

abuse of discretion, the lawyer is accorded the respect appropriate for the exercise of professional judgment.[157]

CONCLUSION

As every new law student quickly discovers, legal reasoning is not susceptible to the same formalistic structures of logic that may be applied in such disciplines as mathematics or the natural sciences. Instead, legal arguments and conclusions are necessarily derived through complex modes of interpretation, often based in potentially imprecise factors including textual analysis, reasoning by analogy, and policy considerations.[158]

Accordingly, a central characteristic of the adversary system is the process through which advocates attempt to procure a favorable disposition through the quality of the arguments they present, not merely as a result of the asserted virtue of their conclusions. Likewise, judicial rulings in hard cases are routinely accompanied by opinions articulating the reasoning behind the decisions. Indeed, the emphasis on the logic behind a conclusion is so central to our legal system that a number of courts accord precedential value only to "published" opinions, in which the reasoning underlying the ruling has been fully delineated.[159]

> whether Dworkin's model of judging (and by analogy Simon's model of legal ethics) adequately handles the case in which two legal actors disagree in good faith about the interpretation of the scheme of justificatory principles underlying the practice of judging or lawyering.

Id. at 74.

157. *Cf.* Stier, *supra* note 20, at 561 ("Since both the decision maker and any subsequent evaluator of that decision will be examining reasons for action, not only behavioral descriptions, explanations and predictions, the centrality of the actor's point of view is particularly important.").

158. *See supra* note 156 and accompanying text.

159. In *Hart v. Massanari*, 266 F.3d 1155 (9th Cir. 2001), Judge Alex Kozinski provided an extensive explanation for the rationale behind this policy:

> Writing a precedential opinion . . . involves much more than deciding who wins and who loses in a particular case. It is a solemn judicial act that sets the course of the law for hundreds or thousands of litigants and potential litigants . . . . An unpublished disposition is, more or less, a letter from the court to parties familiar with the facts, announcing the result and the essential rationale of the court's decision . . . . [A]lthough three judges might agree on the outcome of the case before them, they might not agree on the precise reasoning or the rule to be applied to future cases.

Id. at 1177-78. *See also* Alex Kozinski & Stephen Reinhardt, *Please Don't Cite This! Why We Don't Allow Citation to Unpublished Dispositions*, CAL. LAW., June 2000, at 44.

For additional scholarly treatment of the issue, from the perspectives of judges, law professors, practitioners, and law students, see, e.g., Danny J. Boggs & Brian P. Brooks, *Unpublished Opinions & The Nature of Precedent*, 4 GREEN BAG 2d 17 (2000); Brian P. Brooks, *Publishing Unpublished Opinions: A Review of the Federal Appendix*, 5 GREEN BAG 2d 259 (2002); Kenneth Anthony Laretto, *Precedent, Judicial Power, and the Constitutionality of "No-Citation" Rules in the*

It would seem particularly suitable for ethical decisionmaking similarly to be governed by a model that accepts and appreciates various approaches to inherently complex questions. Nevertheless, as Professor Simon and a number of other scholars have observed, the prevailing model of legal ethics and ethical discretion stands out as a stark exception to American law's development of a modern jurisprudence acknowledging and embracing the complex nature of legal decisionmaking.[160] Anomalously, the area of ethical decisionmaking, in which lawyers must confront some of the most difficult personal and professional issues they are likely to encounter, is also one area in which the American legal system has remained resistant to advancing beyond a structure of mechanical decisionmaking.

In response to both the prevailing model of legal ethics and alternative models proposed by leading ethics scholars, the Deliberative Model aims to offer a framework through which lawyers are both encouraged and, in many cases, required to engage in the kind of complex legal reasoning that is expected of lawyers in virtually all other areas of law. In so doing, the Deliberative Model likewise relies upon both substantive and procedural mechanisms already prevalent in the American legal system. Moreover, through its emphasis on the quality—rather than the outcome—of ethical deliberation, the Deliberative Model accounts for the potential variety of appropriate responses to an ethical dilemma. Ultimately, the Deliberative Model requires that lawyers take ethical discretion seriously, imposing on lawyers an obligation to exercise their discretion through ethical decisions that are the product of articulable and justifiable ethical deliberation.

Federal Courts of Appeals, 54 STAN. L. REV. 1037 (2002); Boyce F. Martin, Jr., *Judges on Judging: In Defense of Unpublished Opinions*, 60 OHIO ST. L.J. 177 (1999); Johanna S. Schiavoni, Comment, *Who's Afraid of Precedent?: The Debate Over the Precedential Value of Unpublished Opinions*, 49 UCLA L. REV. 1859 (2002); Suzanne O. Snowden, *That's My Holding and I'm Not Sticking to It!": Court Rules that Deprive Unpublished Opinions of Precedential Authority Distort the Common Law*, 79 WASH. U. L.Q. 1253 (2001); David S. Tatel, *Some Thoughts on Unpublished Decisions*, 64 GEO. WASH. L. REV. 815 (1996); Carl Tobias, *Anastasoff, Unpublished Opinions, and Federal Appellate Justice*, 25 HARV. J.L. & PUB. POL'Y 1171 (2002); Lance A. Wade, *Honda Meets Anastasoff: The Procedural Due Process Argument Against Rules Prohibiting Citation to Unpublished Judicial Decisions*, 42 B.C. L. REV. 695 (2001).

160. *See supra* notes 41-44 and accompanying text.

FUNDAMENTAL PRINCIPLES FOR CLASS ACTION GOVERNANCE

ALEXANDRA LAHAV*

TABLE OF CONTENTS

INTRODUCTION

Class actions face a crisis of governance. The form of governance provided by Rule 23, governance by representative parties, is both vague in theory and ignored in practice.[1] Instead, by a combination of procedural rules, judicial interpretation and common practice, the class is governed by attorneys with limited judicial oversight. This regime neither reflects the basic insight that the

---

* Teaching Fellow, Stanford Law School. J.D., Harvard Law School. Thank you to Barbara Babcock, Amanda Frost, Pamela Karlan, Mark Kelman, Nick Lundgren, Martha Minow, Henry Monaghan, Alan Morrison, and Allan Stein for valuable comments on previous drafts of this article. I am also grateful to Janet Alexander, Richard Freer, Barbara Fried, Tom Grey, Joseph Grundfest, Deborah Hensler and Paul Lomio for their time and insights.

1. *See* FED. R. CIV. P. 23(a)(4) (requiring representative parties who will "fairly and adequately protect the interests of the class" as a prerequisite to bringing a class action). My discussion of Rule 23 addresses the Rule as amended in December 2003.

class and attorney do not have a traditional attorney-client relationship nor performs the task of transforming the inchoate collectivity of the class into an organization that protects and is responsive to the interests of class members.[2] This Article proposes an alternative regime of governance for 23(b)(3) small claims class actions[3] that accomplishes both these things, based on four fundamental principles: mandatory disclosure of material information, an actively

---

2. As a recent RAND study observed, procedural rules "provide only a weak bulwark against self-dealing and collusion." DEBORAH R. HENSLER ET AL., CLASS ACTION DILEMMAS: PURSUING PUBLIC GOALS FOR PRIVATE GAIN 120 (RAND 2000).

3. This Article focuses on small claims class actions, sometimes referred to as "negative value claims" because litigating the claim will result in a loss to the claim-holder. These are the quintessential class actions and encompass mostly consumer class actions. I define these as claims with a predicted individual award of under $10,000. This Article does not address mass tort class actions, or aggregative litigation, where there are often substantial money damages sufficient to justify individual suits as well as large differentials in potential recovery, punitive damages class actions, or non-opt out class actions brought under Rules 23(b)(1) and 23(b)(2), although the principles articulated here may also be applicable to these cases. I have chosen to address the quintessential class action, which, according to the Supreme Court and other commentators, is the least problematic. Consumer class actions constitute one third of the class actions brought, and thus are a significant presence in the class action world. *See* HENSLER ET AL., *supra* note 2, at 49-123 (describing that in 1995-96, class actions were mostly damages 23(b)(3) class actions, not civil rights or social reform, and estimating that one third of class actions against businesses are consumer class actions not including securities class actions). Many scholars consider consumer damage class actions to be prototypical. *See* Jules Coleman & Charles Silver, *Justice in Settlements*, 4 PHIL. & LAW 102, 122 (Autumn 1986) (describing consumer class actions with small recoveries as "'paradigm class actions'" because no one would bring them if the class action device was unavailable); HENSLER ET AL., *supra* note 2, at 68 (describing actions for money damages as the "traditional paradigm" of class actions). One reason these cases are thought paradigmatic is that they are considered to be a more straightforward use of the class action mechanism than mass tort actions. For example, in *Amchem Products, Inc. v. Windsor*, the Court opined in passing that the "predominance test," so difficult to meet in the mass tort context, is "readily met in certain cases alleging consumer or securities fraud or violations of the antitrust laws." 521 U.S. 591, 625 (1997). *See also* Richard A. Nagareda, *Autonomy, Peace, and Put Options in the Mass Tort Class Action*, 115 HARV. L. REV. 747, 750 (2002) (observing that in the context of securities, antitrust, and consumer litigation "the prospect of effecting a comprehensive peace through a class settlement poses little threat to the autonomy of class members"); John C. Coffee, Jr., *Class Action Accountability: Reconciling Exit, Voice, and Loyalty in Representative Litigation*, 100 COLUM. L. REV. 370, 386 (2000) [hereinafter Coffee, *Accountability*] (arguing that internal conflicts are not typically present in "commercial class actions because . . . in these cases, some objective measure" of damages permits automatic allocation without discretion). Among other things, I hope to show that the same intractable problems faced in mass tort and other more complex contexts are also present in consumer class actions. I hope this Article will illustrate how very problematic small claims or "negative value" class actions are. The law and policy concerning these different types of class actions is so varied that no one system of governance—never mind a single procedural rule—should be applied across the board to all types of class actions.

adversarial process, expertise of decisionmakers, and independence of decisionmakers from influence and self-interest.

In a class action, absent persons, who may or may not want to be part of a lawsuit, receive notice that they are part of a class action or that they are part of a class action settlement.[4] They are denied the power to define the parameters of their group.[5] They do not pick their representatives,[6] and have no power to

---

4. FED. R. CIV. P. 23(c)(2) provides:

> In any class action maintained under subdivision (b)(3), the court shall direct to the members of the class the best notice practicable under the circumstances, including individual notice to all members who can be identified through reasonable effort. The notice shall advise each member that (A) the court will exclude the member from the class if the member so requests by a specified date; (B) the judgment, whether favorable or not, will include all members who do not request exclusion; and (C) any member who does not request exclusion may, if the member desires, enter an appearance through counsel.

5. In many cases, the attorneys will first seek certification of a class with specific parameters, then change the parameters of the class as part of a settlement or in preparation for trial. *See* HENSLER ET AL., *supra* note 2, at 217 (discussing *Selnick v. Sacramento Cable*, where the attorney filed an amended complaint extending the class definition in preparation for trial); *see also* Brian Wolfman & Alan B. Morrison, *Representing the Unrepresented in Class Actions Seeking Monetary Relief*, 71 N.Y.U. L. REV. 439, 444-45 (1996) (discussing the problem of ensuring proportional allocation where there is no "natural class"). Because global peace—or the ability to lay to rest the most possible claims—is of special interest to defendants, the practice of enlarging the scope of the class as part of settlement can smack of collusion.

6. The observation that the class does not pick its attorney and that the class representative is merely a figurehead has become commonplace among scholars. *See, e.g.*, Coffee, *Accountability*, *supra* note 3, at 406 ("Commentators have generally agreed that the representative in a class action is more a figurehead than an actual decision-maker."); John C. Coffee, Jr., *The Regulation of Entrepreneurial Litigation: Balancing Fairness and Efficiency in the Large Class Action*, 54 U. CHI. L. REV. 877, 877 (1987) [hereinafter Coffee, *Entrepreneurial Litigation*] (referring to client control of litigation decisions in the class action as a "noble myth"); Samuel Issacharoff, *Class Action Conflicts*, 30 U.C. DAVIS L. REV. 805, 805 (1997) (noting that courts determine adequacy of representation based on "representations of counsel who have little if any connection to the parties to be bound"); Henry Paul Monaghan, *Antisuit Injunctions and Preclusion Against Absent Nonresident Class Members*, 98 COLUM. L. REV. 1148, 1149 n.1 (1998) ("Once enmeshed in a class action, class members cannot shape their own claims, and their individual rights to participate in the class proceeding are quite limited."). The acceptance of the proposition that the actual relationship between the class and the class counsel is irrelevant is illustrated by the proposed Rule 23(g), which provides a series of factors for court approval of class counsel: the one factor conspicuously absent from the list is client approval. *See also* Stewart v. Gen. Motors Corp., 756 F.2d 1285, 1295 n.5 (7th Cir. 1985) (quoting Alpine Pharmacy, Inc. v. Chas. Pfizer & Co., 481 F.2d 1045 (2d Cir. 1973)).

> One accepting employment as counsel in a class action does not become a class representative through simple operation of the private enterprise system. Rather, both the class determination and designation of counsel as class representative come through

dismiss those representatives.[7] Instead, as a practical matter, class counsel defines the group membership, manages the litigation, makes unilateral strategic decisions, oversees the accrual of fees and costs, and shapes the outcome of a mysterious process class members neither launched nor agreed to resolve.[8] It is not surprising, then, that class actions have been criticized as mechanisms by which attorneys exploit disunited class members and garner enormous fees, while claimants recover minimal damages.[9] Nor is it surprising that 47% of Americans believe that consumer class actions benefit plaintiffs' lawyers most and 20%

---

judicial determinations, and the attorney so benefited serves in something of a position of public trust.

Id. at 1050. There are, however, exceptions to this. Empirical studies have found that class representatives sometimes seek out attorneys, rather than the other way around. *See* HENSLER ET AL., *supra* note 2 (describing examples of different relationships between class counsel and class representatives). Institutional investors may act as lead counsel in securities class actions. 15 U.S.C.A. § 78u-4 (West 2003). *Cf.* Berger v. Compaq Computer Corp., 257 F.3d 475, 483 (5th Cir. 2001) (holding that "PSLRA's requirement that securities class actions be managed by active, able class representatives who are informed and can demonstrate they are directing the litigation" raises the standard adequacy threshold in Rule 23). Thus, the class representative is a representative in the formal sense, that is, that he or she theoretically has the authority to act on behalf of the class, but no conception of popular sovereignty or accountability is imbedded in the mechanisms of representation.

7. This is because after certification, class counsel's ethical obligations run to the class, not to the individual class member. *See In re* Gen. Motors Corp. Pick-Up Truck Fuel Tank Prods. Liab. Litig., 55 F.3d 768, 801 (3d Cir. 1995) ("Beyond their ethical obligations to their clients, class attorneys, purporting to represent a class, also owe the entire class a fiduciary duty once the class complaint is filed."); Jonathan R. Macey & Geoffrey P. Miller, *The Plaintiffs' Attorney's Role in Class Action and Derivative Litigation: Economic Analysis and Recommendations for Reform*, 58 U. CHI. L. REV. 1, 20 (1991) (noting that "the attorney must act for the benefit of the class as a whole and therefore is not obliged to follow the unilateral wishes of any individual class member").

8. Although there may be a presumed requirement that the attorneys consult with the class representatives, if these representatives have been hand-picked by the attorney, such consultation may have little meaning. In any event, the Rules create no obligation for class counsel to consult with the class or answer to class representatives prior to notifying the class or presenting a settlement to the court for approval pursuant to Federal Rule of Civil Procedure 23(e).

9. *See, e.g.*, Jesse J. Holland, *Official: Put Class Actions in U.S. Court*, AP ONLINE, May 15, 2003, *available at* 2003 WL 55372125 (stating that Justice Department spokesperson asserted that trial lawyers benefit more from class actions than plaintiffs they represent); *see also* Deborah R. Hensler, *Revisiting the Monster: New Myths and Realities of Class Action and Other Large Scale Litigation*, 11 DUKE J. COMP. & INT'L L. 179, 180 (2001) [hereinafter Hensler, *Revisiting the Monster*] (noting popular belief that "monsters are loose in the land"); Susan P. Koniak & George M. Cohen, *Under Cloak of Settlement*, 82 VA. L. REV. 1051 (1996) (arguing that attorneys often abuse their position as class counsel and take advantage of class members in order to increase their fees); Deborah R. Hensler & Thomas D. Rowe, Jr., *Beyond "It Just Ain't Worth It": Alternative Strategies for Damage Class Action Reform*, 64 LAW & CONTEMP. PROBS. 137, 137-39 (2001) [hereinafter Hensler & Rowe, *Beyond "It Just Ain't Worth It"*] (collecting similar criticisms).

believe that they benefit attorneys for defendants most, while only 9% believe that class actions benefit plaintiffs most.[10]

A few procedural protections purport to lend legitimacy to this system by enlisting judicial oversight to prevent gross abuses of the class action. In certifying a class, the court reviews the adequacy of the representatives and counsel.[11] After the class is constituted and settlement has been reached, class members are given opportunities to opt out or to speak out at a fairness hearing.[12] Nevertheless, an absent class member, not represented by counsel and armed with minimal information, has a very limited ability to take advantage of these mechanisms.[13] In the best case, class members' interests are represented by responsible objectors. But for the most part, the class is left to rely on the expertise of class counsel and the wisdom of judges to guarantee fairness.[14] Judges approve the class parameters, decide whether or not a settlement is fair, and push the litigation in the direction of settlement or adjudication.[15] Notwithstanding the judge's role in overseeing class actions, fair results are far from guaranteed.

The dispute over whether the benefits of the small claims class action

---

10. *Opinions on Class-Action Lawsuits*, USA TODAY, Mar. 24, 2003, at B1, *available at* 2003 WL 5307581.

11. *See* FED. R. CIV. P. 23(a) (setting forth the requirements for class certification). Certification is no easy feat for plaintiff's counsel to pull off, but nevertheless, meeting the criteria of 23(a) provides no guarantee of good governance.

12. *See* Fed. R. Civ. P. 23(c)(2) (requiring notice and opportunity to opt out at certification to (b)(3) class members); 23(e) (class action may not be dismissed or compromised without judicial approval); 23(e)(1)(B) (requiring notice of settlement); 23(e)(1)(C) (court may approve settlement upon finding that it is "fair, reasonable, and adequate"); 23(e)(3) (opportunity to opt out at settlement is at court's discretion).

13. One Federal Judicial Center study found that only seven to fourteen percent of fairness hearings were attended by class members who were not objectors or named plaintiffs and that in forty-two to sixty-four percent of fairness hearings there were no objectors at all. Koniak & Cohen, *supra* note 9, at 1105-06.

14. *See* William B. Rubenstein, *Divided We Litigate: Addressing Disputes Among Group Members and Lawyers in Civil Rights Campaigns*, 106 YALE L. J. 1623, 1663-68 (1997) (discussing the benefits and drawbacks of an expertise-based model of group governance). *Cf.* Coffee, *Accountability*, *supra* note 3, at 406 (characterizing the possibilities of voice under the current Rule 23 regime as a choice "between the uninformed democracy of class members versus the often self-interested professionalism of plaintiffs' attorneys").

15. *See* FED. R. CIV. P. 23(c) (power of court to determine when class action may be maintained); 23(c)(3) (to determine membership in class); 23(c)(4) (to require sub-classification); 23(d) (to issue procedural orders); 23(e) (to approve dismissal or compromise). *See generally* Judith Resnik, *Whose Judgment? Vacating Judgments, Preferences for Settlement, and the Role of Adjudication at the Close of the Twentieth Century*, 41 U.C.L.A. L. REV. 1471 (1994) (arguing that litigants do not expect to go to trial and that judges, by using rules, doctrine, and practices try to get them to settle).

outweigh its costs is longstanding, much discussed, and unresolved.[16] Among other things, class actions solve the collective action problems faced by individuals with claims too small to be economically adjudicated individually, and address certain small private wrongs with substantial public effects, especially in the absence of governmental intervention.[17] Thus, class actions serve a regulatory function.[18]

There are two substantive justifications for permitting groups to litigate through the class action mechanism: compensation and deterrence.[19] Because the

---

16. "For all our effort, we do not know whether this is a good or a bad thing. The great big question is whether the social utility of the large class action outweighs the limited benefits to individuals, the aroma of gross profiteering, and the transactional costs to the court system." John Frank, 1966 Civil Rules Advisory Committee Member in a memorandum to the chair of the 1995 Committee (quoted in HENSLER ET AL., *supra* note 2, at 401). *See also* Hensler & Rowe, *Beyond "It Just Ain't Worth It," supra* note 9, at 137-39 (describing ongoing debate over merits of damages class actions); Hensler, *Revisiting the Monster*, *supra* note 9, at 196 (arguing whether class actions are directed at doing social good or ill "depends substantially on how well judges control the litigation process"). *But see* Elizabeth J. Cabraser, *Enforcing the Social Compact Through Representative Litigation*, 33 CONN. L. REV. 1239 (2001) (arguing that class action litigation is a beneficial social good).

17. In situations where claims are small, most of the individual class members "would have no realistic day in court if a class action were not available." Phillips Petroleum Co. v. Shutts, 472 U.S. 797, 809 (1985). *See also* Deposit Guar. Nat'l Bank v. Roper, 445 U.S. 326, 338 n.9 (1980) (describing negative value class action claims); Coleman & Silver, *supra* note 3, at 121 n.60 ("A rational person will refuse to sue on a claim when the opportunity cost of a lawsuit exceeds its expected return (the payoff given that the lawsuit may be won, lost or settled times the probability of each outcome)."). As Arthur Levitt, former Chairman of the SEC, observed, "private litigation is 'the primary vehicle for compensating defrauded investors.'" HENSLER ET AL., *supra* note 2, at 69. The class action mechanism seems to have competing objectives: administrative efficiency and enabling litigation in pursuit of larger social goals. *See id.* at 49. Hensler points out that the "distinction in the public debate between the *efficiency* and *enabling* goals of class actions for money damages is illusory. In practice, any change in court processes that provides more efficient means for litigating is likely to enable more litigation." *Id.* In recent years, there has been an increase of involvement of state attorneys general in consumer issues. This development is an interesting one and perhaps should affect the way policy-makers look at the regulatory role of consumer class actions. On the other hand, this development may be too personality-driven to spur a change in policy towards class actions.

18. "[M]any see class actions as a powerful regulatory enforcement tool and view those who bring them as 'private attorneys general.'" Hensler, *Revisiting the Monster*, *supra* note 9, at 182-83.

19. Courts have interpreted the class action mechanism as a way to solve the collective action problem of harms that are too small for individuals to prosecute. At the core of this goal is compensation for those individuals. Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 617 (1997) ("the very core of the class action mechanism is to overcome the problem that small recoveries do not provide the incentive for any individual to bring a solo action prosecuting his or her rights. A class action solves this problem by aggregating the relatively paltry potential recoveries into

aggregation of these small claims has a substantial effect on the defendant, deterrence may have a larger influence than compensation.[20] It may be that the smaller the likely compensation to the individual, the more important the deterrence justification becomes.[21] To the extent that the class action's primary goal is compensation, the objective of its governance scheme should be to maximize claimholder value.[22] To the extent that the primary goal is deterrence, the governance regime should be concerned that the amount extracted from defendants is sufficient to further that goal and is properly distributed among class members, the public, and class counsel. This Article assumes that the

---

something worth someone's (usually an attorney's) labor." (quoting Mace v. Van Ru Credit Corp. 109 F.3d 338, 344 (7th Cir. 1997))); 5 JAMES WM. MOORE ET AL., MOORE'S FEDERAL PRACTICE § 23.02 (3d ed. 1998). Torts scholars have moved towards the view that tort norms, such as deterrence, are a more appropriate way to view class actions. *See* David Rosenberg, *Individual Justice and Collectivizing Risk-Based Claims in Mass-Exposure Cases*, 71 N.Y.U. L. REV. 210, 214 (1996) (proposing a functional approach to mass tort actions, viewing them from point of view of tort policy of minimizing costs of accidents and maximizing deterrence). Scholars recognize a larger social goal for securities class actions. *See* Janet Cooper Alexander, *Rethinking Damages in Securities Class Actions*, 48 STAN. L. REV. 1487 (1996) (observing that the purpose of securities class actions is to protect the public interest and the integrity of capital markets). These justifications necessarily exclude frivolous class actions brought by attorneys intent on obtaining fees to "blackmail" defendants. *See, e.g.*, Monte Morin, *Lawyers Who Sue to Settle*, L.A. TIMES, Oct. 6, 2002, at A1 (describing attorneys who file suits with the intent to settle for attorneys fees). These types of "blackmail" class actions are not an internal governance problem of the type addressed here, but are an important issue in class action policy generally. For mechanisms to screen class actions, see Hensler & Rowe, *Beyond "It Just Ain't Worth It," supra* note 9. For a discussion of blackmail class actions and suggestions for reform, see Bruce Hay & David Rosenberg, *"Sweetheart" and "Blackmail" Settlements in Class Actions: Reality and Remedy*, 75 NOTRE DAME L. REV. 1377 (2000).

20. A pure deterrence perspective does not provide a basis for distributing the proceeds among the class members, but only requires that they be disgorged from the defendant. *Cf.* Coleman & Silver, *supra* note 3, at 137 (arguing that the principle of compensation does not require that the settlement fund be paid out to class members). It nevertheless seems fundamentally unfair that compensation should mostly go to attorneys at the expense of wronged class members, even if the wrongs themselves are small. When claimants are difficult to find, a charitable fund is one solution. *See In re* Mex. Money Transfer Litig., 164 F. Supp. 2d 1002 (N.D. Ill. 2000) (approving settlement that included *cy pres* fund where claimants were illegal immigrants and thus difficult to find), *aff'd*, 267 F.3d 743 (7th Cir. 2001).

21. *See* David L. Shapiro, *Class Actions: The Class as Party and Client*, 73 NOTRE DAME L. REV. 913, 923-31 (1998) (arguing that the main purpose of small claims class actions is deterrence).

22. This is with the caveat that the amounts meted out in settlement should reflect the value of the claim at litigation. *See* Janet Cooper Alexander, *Do the Merits Matter? A Study of Settlements in Securities Class Actions*, 43 STAN. L. REV. 497, 566 (1991) (showing that in many securities litigation settlements the amounts paid out cohere to a "going rate" of settlement rather than an amount reflecting the strength of plaintiffs' claims).

regulatory device of the class action attempts to achieve deterrence through individual compensation. Whether this device is effective in achieving these linked goals—and whether they need to be linked at all—is an open question.[23]

One might ask why we should care about class action governance at all. Scholars viewing class actions from an economic perspective generally view the central problem in class actions to be an agent-principal problem, and have a correspondingly narrow view of the governance regime required to resolve it.[24] Other scholars assume that participation values have a place in class actions, but without providing the conceptual underpinning for this assumption beyond the representation requirement in Rule 23.[25] The due process requirements of notice, hearing, and opportunity to opt out in 23(b)(3) class actions, combined with the Rule 23(a)(4) representation requirements,[26] indicate that the rule makers saw some autonomy value for individual claimants in class action litigation, but the strength of that value is uncertain and the reasons for it are not clarified in the case law.[27] The smaller the individual claim at issue, as in consumer class

---

23. Arguably the smaller the individual claim the lower the autonomy value associated with that claim. *See, e.g.*, Gammon v. GC Servs. Ltd. P'ship, 162 F.R.D. 313, 321-22 (N.D. Ill. 1995) (holding that *cy pres* distribution of damage award is appropriate where class of four million had claim valued at thirteen cents each). In such cases, the compensation goal may be a barrier to extracting the correct amount from the defendant and thus a barrier to deterrence.

24. The underlying assumption of economic analysis of this type appears to be that compensation is the central goal of consumer class actions. *See, e.g.*, John C. Coffee, Jr., *Understanding the Plaintiff's Attorney: The Implications of Economic Theory for Private Enforcement of Law Through Class and Derivative Actions*, 86 COLUM. L. REV. 669 (1986) [hereinafter Coffee, *Understanding the Plaintiff's Attorney*]; Macey & Miller, *supra* note 7. *See also* Judith Resnik et al., *Individuals Within the Aggregate: Relationships, Representation, and Fees*, 71 N.Y.U. L. REV. 296, 374 (1996) (describing how economic analysis of litigation has been based on the agent-principal problem).

25. *See, e.g.*, Owen M. Fiss, *The Political Theory of the Class Action*, 53 WASH. & LEE L. REV. 21, 25 (1996) (criticizing the self-selected representative in class actions as an "anomalous" form of governance); Shapiro, *supra* note 21, at 958-59 (advocating for fuller representation of class members as a group).

26. Federal Rule of Civil Procedure 23(a)(4) requires that the class be "adequately" represented in order to be certified as a class.

27. For an enlightening discussion on the basis for and critique of this individual choice rationale, see Robert G. Bone, *Rethinking the "Day in Court" Ideal and Nonparty Preclusion*, 67 N.Y.U. L. REV. 193 (1992). *See also* Patrick Woolley, *Rethinking the Adequacy of Adequate Representation*, 75 TEX. L. REV. 571 (1997) (arguing that due process requires notice, the right to be heard, and opt out in the class action context and that these are inadequately met under the current regime). Autonomy values, the traditional justification for due process protections and good governance, are less powerful in this context. *See* Frank Michelman, *Formal and Associational Aims in Procedural Due Process*, DUE PROCESS: NOMOS XVIII 126-28, 154 n.4 (J. Pennock & J. Chapman eds., 1977) ("[R]espect for individual dignity, autonomy, and self-expression demands that those with rights directly at risk have adequate means of registering their concerns."). Since the claim is small, some might argue, as long as there are mechanisms to avoid fraud or self-dealing

actions, the more limited the autonomy values that can be placed on that individual claim.

Despite small individual recoveries, the internal governance of small claims class actions deserves attention.[28] Both the regulatory function of the class action mechanism and the deterrence justification are promoted by a robust governance regime. Furthermore, a strong governance regime will benefit class members by increasing agent monitoring to prevent exploitation. Good governance will improve outcomes in class actions by multiplying the sources of protection against agency problems and by giving advocates of alternative solutions an opportunity to be heard.[29] Furthermore, because consumer class actions seek to

---

in payouts, the internal governance structure does not matter because the small size of the claims renders autonomy values relatively weak. But autonomy values are not the only values for support of participatory governance. Viewed collectively, the size of these cases is significant. Communitarian values may be a more fitting place to look for class action governance. *Cf.* MICHAEL SANDEL, LIBERALISM AND THE LIMITS OF JUSTICE 173-74, 183 (contrasting "liberalism" with a "constitutive conception of community"). Although it is difficult to describe the consumer class as a "community" in the way that community is typically defined in political science literature, it nevertheless is possible to see the consumer class as a group with a public-oriented interest beyond their individual interest in compensation. *Cf.* Frank I. Michelman, *Conceptions of Democracy in American Constitutional Argument: Voting Rights*, 41 FLA. L. REV. 443, 445 (1989) (comparing liberal, individually oriented conceptions with the republican conception of an autonomous public interest).

28. The reasons for the importance of the internal governance structure of class settlements presented here are admittedly incomplete. I hope to develop a more rigorous analysis of the organizational theory of the class action and to explore more fully the extent to which norms other than litigant autonomy can serve as a basis for a robust governance regime in the future.

29. Some alternative options include the use of *cy pres* funds in addition to or instead of compensation directly to class members. *See* 3 NEWBERG ON CLASS ACTIONS § 10:17 (4th ed. 2003) (describing use of *cy pres* distributions of class damages); Kerry Barnett, *Equitable Trusts: An Effective Remedy in Consumer Class Actions*, 96 YALE L. J. 1591 (1987). Other alternatives may be the provision of repairs or retrofitting. *See In re* Ford Motor Co. Bronco II Prod. Liab. Litig., 1995 U.S. Dist. Lexis 3507 (E.D. La. Mar. 15, 1995) (noting objector suggestion that retrofitting would significantly improve stability and safety of vehicle at issue in litigation). Although there is no empirical evidence that good governance, or that this governance proposal in particular, will yield better substantive results to class action suits, what seems clear is that the current system of attorney control does not yield excellent substantive results. Furthermore, vociferous criticism of plaintiffs' side class action attorneys indicates the current system of governance has been difficult to justify on a results basis. This utilitarian argument has been one basis for liberal political philosophy. *See* JOHN LOCKE, SECOND TREATISE OF GOVERNMENT 56 (Richard H. Cox, ed. 1985).

> [A]s if when men quitting a state of nature entered into society, they agreed that all of them but one should be under the restraints of laws, but that he should still retain all the liberty of the state of nature, increased with power, and made licentious with impunity. This is to think that men are so foolish that they take care to avoid what mischiefs may be done to them by *polecats* or *foxes*, but are content, nay think it safety, to be devoured

deter wrongdoing, they are a public good. Thus, the inclusion of deliberative process, accountability and responsiveness, to claimants and the public, in class action governance is justified for the same reasons notice and comment provisions are integral to administrative and regulatory law.[30] Finally, good governance would give class action architects greater public legitimacy because they can point to a process that resembles other legitimate processes for reaching decisions in our society, in contrast to the current regime that increases the perception that attorneys are exploiting class actions unchecked.[31]

In pursuit of this goal, this Article looks at the class as a group in need of governance and proposes a comprehensive set of principles through which to judge the efficacy of class action mechanisms. This Article focuses on governance in the settlement phase because it is one of the most contentious moments in class action litigation. Part I illustrates the problems in the current regime of class action governance for small claims class actions. Part II analyzes and critiques three solutions to these problems that have been proposed over the last twenty years: market mechanisms, democracy-based solutions, and judicial administration. Part III proposes a comprehensive model of governance based on four fundamental principles: mandatory disclosure, an actively adversarial process, expertise, and independence of decision-making.

---

by *lions*.

Id.

30. The Administrative Procedure Act, for example, provides for notice and comment on rules and thus indicates support for participation values in the administrative state. *See also* NLRB v. Wyman-Gordon Co., 394 U.S. 759, 764-65 (1969) (opining public participation in administrative rulemaking proceedings ensures that regulation is responsive to the needs of those regulated). Both the defendants, who are represented directly in the proceedings, and class members, have concrete interests in the outcomes and distributive effects of class actions. The public has a more attenuated interest in these outcomes. In some circles, scholars are cynical about the relationship between administrative and regulatory agencies and notions of popular sovereignty. Although there is insufficient room to develop this issue further, there is significant work to be done on the relationship between democratic values in regulatory litigation and the regulatory state.

31. The pervasive criticisms against class actions—which fall mostly in the lap of plaintiffs' side class action attorneys—are damaging to the legal profession and to the regulatory purpose of the class action mechanisms. Commentators who oppose class action litigation often point to lawyer control as the reason for eliminating class actions. *See, e.g.*, WALTER OLSON, THE RULE OF LAWYERS: HOW THE NEW LITIGATION ELITE THREATENS AMERICA'S RULE OF LAW (2003). Studies and opinion polls have found that the public believes that lawyers are greedier and more dishonest than other professionals. *See* Lynn A. Baker & Charles Silver, *Introduction: Civil Justice Fact and Fiction*, 80 TEX. L. REV. 1537, 1539 (2002) (describing studies). The bad press concerning class action litigation has not helped this image. Good governance will lend legitimacy to the class action process and by extension to the legal profession charged with overseeing these cases.

I. PROBLEMS IN CLASS ACTION GOVERNANCE

A. A Misapplication of the Traditional Adjudication Model

Class action governance problems are rooted in the misapplication of the traditional model of adjudication to the group context. While it may seem obvious that a class action is more than simply a suit with multiple complainants, such has been the power of the traditional model of adjudication that very few procedural mechanisms were deemed necessary to ensure fair and efficient class governance.[32] In reality, the contrast between the traditional model and class action litigation is sharp.

In the traditional model of adjudication, the client is a voluntary and active participant in the lawsuit, dictating its course.[33] The attorney is bound by specific, enforceable ethical duties to zealously advocate on behalf of the client's interests, to keep the client informed, and to consult the client at every critical juncture in the litigation. The client's leverage over his agent is ensured by his power to hire or fire the attorney at will. Finally, the judge is a neutral arbitrator of the dispute.[34]

Class actions set the traditional model of adjudication on its head.[35] The class action effectively herds absentee plaintiffs into a lawsuit without their consent and often without their knowledge in part because the right to opt out is illusory. Class counsel fills the resulting power vacuum by proposing the parameters of the class, recruiting named representatives, and making every

---

32. Class members are protected by a determination that the representation is adequate, notice, a hearing, and the opportunity to opt out. *See* FED. R. CIV. P. 23. On the history of the development of Rule 23, see HENSLER ET AL., *supra* note 2, at 9-37.

33. *See* Abram Chayes, *The Role of the Judge in Public Law Litigation*, 89 HARV. L. REV. 1281, 1282 (1976) (characterizing the traditional model of civil litigation as having the following defining features: (i) it is bipolar, (ii) litigation is retrospective, (iii) right and remedy are interdependent, (iv) the lawsuit is a self-contained episode, (v) the process is party initiated and party controlled). This traditional model is an ideal type even when applied to individual litigation, but has powerful resonance that has been particularly damaging in the class action context. *See generally* Lon L. Fuller, *The Forms and Limits of Adjudication*, 92 HARV. L. REV. 353 (1978) (basing legitimacy of adjudication on highly individualized participation).

34. *See generally* Marvin E. Frankel, *The Search for Truth: An Umpireal View*, 123 U. PA. L. REV. 1031 (1975) (discussing the role of the judge in the search for truth).

35. *See* Samuel Issacharoff, *Preclusion, Due Process and the Right to Opt Out of Class Actions*, 77 NOTRE DAME L. REV. 1057, 1058 (2002) ("A class action is simply, when all else is stripped away, a state-created procedural device for extinguishing claims of individuals held at quite a distance from the 'day in court' ideal of Anglo-American jurisprudence."). Some scholars have made proposals from the point of view that the role of class action governance is to aggregate individual preferences with a strong tilt towards autonomy values. *See, e.g.*, Coffee, *Accountability*, *supra* note 3, at 379-80 (criticizing the "entity theory" as "a legal fiction" that "provides a justification for the attorney making judgments in the 'best interests' of the class and allocations among class members in precisely the manner *Amchem* seemed to forbid").

important decision, including whether to accept or reject proposed settlements.[36] Class counsel are not required to attempt to identify or build majority support for a settlement or survey class members to determine their interests. Far from being a neutral or disinterested arbiter, the judge must often assume an active role in facilitating settlement, in the process becoming interested in its approval.[37]

Despite these differences, few procedural modifications have been made to the class action litigation process to accommodate the collective nature of class action litigation. Instead of dictating the course of the suit, class members' interests are pursued by a self-appointed or court-appointed "governor" who decides what is in their best interest, often without input from his clients. As Owen Fiss points out, "self-appointment is an anomalous form of representation, only justified, if at all, by the most exceptional circumstances."[38] This form of government resolves the problem of organizing a group into a "client" by reducing the group to a few representatives of dubious legitimacy. The results have been occasional instances of egregious corruption on the part of attorneys who take advantage of class members[39] and a perception that consumer class actions are not a public good, but a money making scheme for unscrupulous lawyers.

B. The Simultaneous Expansion and Disintegration of the Class Action

Two forces work to expand the numbers of claimants included in class actions. First, preclusion rules, anti-suit injunctions, and removal statutes

---

36. Coffee, for example, likens class counsel to a joint-venturer with the class, who has a financial stake in the outcome of the litigation because of the considerable investment a class action requires. He argues that the "leading cause of 'cheap settlements' may not be collusion between class counsel and defendants . . . but rather a basic differential in the level of risk aversion" because class counsel, with its significant financial investment in the action, is more likely to be risk averse than the class. Coffee, *Accountability*, *supra* note 3, at 390-93. *See also In re* Cendant Corp. Litig., 264 F.3d 201, 254-55 (3d Cir. 2001) (explaining that "[b]ecause of this conflict (and because 'the class' cannot counteract its effects via counsel selection, retention, and monitoring), an agent must be located to oversee the relationship between the class and their lawyers. Traditionally, that agent has been the court."). While not quite collusion, this perverse incentive does look like a type of self-dealing.

37. Owen Fiss discussed this phenomenon in the context of what he perceived to be the bureaucratization of the judiciary, predicting that where the judge serves "as both architect and structural engineer" of an institution (and, analogously, a settlement), he is likely to view challenges to that institution as personal challenges. Owen M. Fiss, *The Social and Political Foundations of Adjudication*, 6 LAW & HUM. BEHAV. 121, 126 (1982).

38. *See* Fiss, *supra* note 25, at 25.

39. *See generally* Koniak & Cohen, *supra* note 9, at 1058-69 (describing litigation in which class members who received no benefit had attorneys fees deducted from their escrow accounts); Wolfman & Morrison, *supra* note 5, at 473-77 (describing coupon settlements that did not benefit classes).

encourage the consolidation of rival actions in a single forum.[40] Second, class counsel has incentives to settle the most comprehensive action first, because only one action, if any, will ultimately stand. Defendants play rival class counsel against each other to settle for the lowest price, a process referred to as a reverse auction.[41] The phenomenon of reverse auctions, or the threat of reverse auctions, pushes class counsel towards sub-optimal settlements and more global settlements to ensure their own compensation.[42] Enterprising counsel may try to bolster their settlement potential by buying out competitors.[43] When competing attorneys know that they will be precluded by a settlement, a buy-out is an attractive alternative.

The trend towards the expansion of small claims class actions over an ever-increasing number of geographically widespread claims renders the internal governance structure of class actions even more important.[44] The proposed Class Action Fairness Act sought to further expand this trend by granting federal diversity jurisdiction to state law-based small claims class actions.[45] Multi-District Litigation panels consolidate multiple class actions to a single district for discovery and settlement purposes.[46] The right to collateral attack has been limited so that settlement in one forum will extinguish claims in other fora, even claims that could not have been brought in the settling forum.[47] Settling parties

---

40. This in part because rival attorneys know ex ante that these procedural tools will be used against them.

41. *See* John C. Coffee, *Class Wars: The Dilemma of the Mass Tort Class Action*, 95 COLUM. L. REV. 1343, 1370 (1995) (describing the reverse auction phenomenon as "a jurisdictional competition among different teams of plaintiffs' attorneys in different actions that involve the same underlying allegations").

42. *See* Coffee, *Accountability*, *supra* note 3, at 390-93 (discussing risk aversion of class counsel).

43. For example, in the Bausch & Lomb class action discussed in HENSLER ET AL., *supra* note 2, at 158, tag along class actions were settled for undisclosed, but reportedly nominal amounts.

44. The numbers of nationwide class actions are reportedly rising. *See id.* at 64-66. There also is a countervailing trend towards localization of class actions. *See, e.g.*, R.J. Reynolds Tobacco Co. v. Engle, 649 So.2d 245 (Fla. Dist. Ct. App. 1996) (denying certification of nationwide class and approving certification of class of Florida citizens in tobacco class action).

45. *See, e.g.*, Class Action Fairness Act of 2003, S.274, 108th Cong. § 4 (2003) (among other things, proposing amendment to 28 U.S.C.A. § 1332 (West 2003) permitting original jurisdiction of class action where the amount of the entire controversy exceeds two million dollars). Such an amendment would abrogate *Snyder v. Harris*, 394 U.S. 332 (1969) and *Zahn v. International Paper Co.*, 414 U.S. 291 (1973).

46. *See* 28 U.S.C. § 1407 (West 2003).

47. *See* Matsushita Elec. Indus. Co. v. Epstein, 516 U.S. 367 (1996) (holding that the Full Faith and Credit Act requires federal courts to give preclusive effect to state judgments even when state court judgment at issue incorporates class action settlement releasing claims solely within jurisdiction of federal courts, so long as they comply with due process). Subsequently, on remand the 9th Circuit held that the Supreme Court's decision determined that the state court judgment comported with due process. *See* Epstein v. Matsushita Elec. Indus. Co., 179 F.3d 641 (9th Cir.),

can obtain an anti-suit injunction preventing class members from bringing suits in any other forum.[48] Or, they may seek a determination that a rival action is barred in the forum where a competing suit was filed.[49]

For example, in the *Mexico Money Transfer Litigation*,[50] a nationwide consumer class action arising out of allegations that money transfer companies overcharged customers by manipulating the exchange rate, a total of eight lawsuits were initially filed.[51] Most of these were dismissed, dropped or stayed by the federal court in the Northern District of Illinois that ultimately decided the case.[52] After this consolidation, National Class Counsel, consisting of the attorneys who had filed class actions in Texas and Illinois, negotiated with the defendants.[53] This process eliminated potential objections *ex ante*.[54]

---

cert. denied, 120 S. Ct. 497 (1999).

> It is now settled that a judgment pursuant to a class settlement can bar later claims based on the allegations underlying the claims in the settled class action. This is true even though the precluded claim was not presented, and could not have been presented, in the class action itself.

In re Prudential Ins. Co. of Am. Sales Practice Litig., 261 F.3d 355, 366 (3rd Cir. 2000). *See also* Marcel Kahan & Linda Silberman, *The Inadequate Search for "Adequacy" in Class Actions: A Critique of Epstein v. MCA, Inc.*, 73 N.Y.U. L. REV. 765 (1998) (criticizing the right to collaterally attack a settlement for inadequacy of representation as excessively broad). The right to collateral attack in state court appears to be expanding somewhat. *See* State v. Homeside Lending, Inc., 826 A.2d 997 (Vt. 2003).

48. *See* Monaghan, *supra* note 6, at 1150, 1151-52 n.13. (discussing *Baker v. General Motors' Corp.*, 118 S. Ct. 657, 665 (1998), which distinguished between enforcement and preclusive effects of antisuit injunctions). As Monaghan points out, the nature of the relationship between the enjoining court and the enjoined court has yet to be worked out.

49. *See* Syngenta Crop Prot. Inc. v. Henson, 123 S. Ct. 366, 379 n.1 (2002).

50. 164 F. Supp. 2d 1002 (N.D. Ill. 2000), *aff'd*, 267 F.3d 743 (7th Cir. 2001), *cert. denied*, Garcia v. W. Union Fin. Servs., Inc., 535 U.S. 1018 (2002).

51. In *In re Mexico Money Transfer Litigation*, two suits were filed in federal court in California. After the federal claims of the initial suits were dismissed, these were refiled in California state court on behalf of a California class with solely California state claims. A third federal suit was filed in California federal court naming a nationwide class. Two class actions were filed in federal court in the Northern District of Illinois. These were subsequently consolidated. Two class actions were filed in federal court in Texas on behalf of a Texas class. Another class action was filed in Texas state court also on behalf of a Texas class. *See* 164 F. Supp. 2d at 1008-09.

52. *Id.*

53. 164 F. Supp. 2d at 1010. The district court noted that the parties' declarations in support of settlement described a year-long, arms-length negotiation and that attorney's fees were only negotiated after a preliminary settlement had been reached.

54. The objectors who were represented by counsel consisted of California class members who alleged that they would have stronger claims under California law. The remaining objectors were not represented by counsel. While it is still possible for the remaining class members to object, many potential committed objectors will opt out.

Concurrent with the expansion of class actions is a disintegration of the courts' power to consolidate and control them. This is a federalism issue as well as one of manageability. Blatant forum shopping is rife in class actions. Class counsel will file national actions in state court in order to gain the leverage of hospitable state courts. Both class counsel and defendants will push class actions towards state court to obtain settlement approval.[55] A settlement rejected in federal court on fairness grounds may be re-filed in state court and approved there.[56] Both class counsel and defendants will pull towards federal court when they believe they can gain more leverage or obtain global settlement in that forum. Objectors and rival class counsel can also use procedural rules to their advantage.[57] Even when a state suit disrupts the settlement already reached and approved in federal court, it is up to the state court to determine whether or not the suit is barred. Collateral attacks in state courts on personal jurisdiction grounds based on denial of due process protections are also gaining momentum.[58] All the while, the courts work to prevent the multiplication of class actions to conserve the efficiency of the class action mechanism and judicial resources. A viable governance regime for class actions must recognize the prevalence of forum shopping and the corollary phenomena of buy-outs of competition by class counsel and defendants.

C. Barriers to Participation

Significant systemic barriers block direct class member participation in class actions. Under the current governance regime, class members have a minimal role in the litigation.[59] Class counsel's role as entrepreneur encourages counsel

---

55. The reasons for state court's apparent deference to settlements are not clear, but anecdotal evidence suggests that this is the case.

56. For example, when the Third Circuit overturned a somewhat notorious settlement in the *In re General Motors Pick-Up Truck Fuel Tank Products Liability Litigation*, the parties re-filed the settlement in Louisiana state court, where it was approved. 134 F.3d 133 (3d Cir. 1998). *See* Public Citizen Litigation Group, *Public Citizen's Involvement in Class Action Settlements* (Feb. 28, 2002) at http://www.citizen.org/litigation/briefs/Class_Action/articles.cfm?ID=552.

57. For example, rival attorneys or objectors can turn to state courts and the All Writs Act cannot furnish removal jurisdiction where a federal court has no original jurisdiction. *See* Syngenta Crop Prot. Inc. v. Henson, 537 U.S. 28 (2002). In *Syngenta*, a suit was settled in a consolidated federal court proceeding with plaintiff's counsel participating. That same counsel then brought a competing class action in state court and argued that the settlement did not preclude the subsequent state court action. The Supreme Court held that the state court would decide whether or not the subsequently filed state court action violated the federal court settlement. *Id.*

58. *See* State v. Homeside Lending, Inc., 826 A.2d 997 (Vt. 2003) (holding that class action settlement did not preclude suit by Vermont residents because Alabama court lacked personal jurisdiction due to defects in due process protections).

59. Although class members have a formal right to participate in the class action, participation is difficult and costly. Structural barriers prevent class members from exercising their right to participate, particularly in the absence of attorney representation.

to keep information from the class in order to retain full control of the negotiation and settlement process. Communication with far-flung class members is expensive, and will likely lead to the discovery of class member preferences or internal class conflicts that class counsel would rather ignore. Were there a requirement that class counsel solicit class member views, or that a certain percentage of the class "opt-in" to the class action by submitting claims, this relationship would change dramatically.

Class members are not repeat players and they cannot create ongoing relationships with class counsel, therefore they lack leverage over their attorney. Because there is no ongoing relationship, even the type of leverage that a consumer has over a manufacturer is lost in the class action context. In a commercial enterprise, a consumer's decision not to buy a product has a direct effect on the bottom line, even if a small one. In the class action context, opt outs will not affect attorney remuneration unless they are so massive as to scuttle the entire settlement. Because of the ongoing relationship between consumers and manufacturers, a decline in services or the product in the commercial enterprise may activate a consumer response to alert the firm to the problem and lead to recovery.[60] There is no such ongoing relationship in the class action.

Structural realities and perverse incentives inhibit participation to an even greater extent in settlement classes. The first notice the class receives of the litigation is also a notice of settlement. If they are dissatisfied with the settlement, they may opt out or protest the settlement in the fairness hearing. The hearing is usually a non-adversarial proceeding because both parties' attorneys want approval of the settlement they have worked hard to formulate.[61] Worse yet, the client—that is, the class—exists at the whim of the defendant because the

---

60. *See* ALBERT HIRSCHMAN, EXIT, VOICE AND LOYALTY: RESPONSES TO DECLINE IN FIRMS, ORGANIZATIONS AND STATES 15 (1970). In order for exit to be an effective means of communicating consumer dissatisfaction, enough individuals must defect to another firm in order to affect the company's bottom line. But unless the company is vigilant, which a firm in decline is unlikely to be, by the time the company understands consumer dissatisfaction, the decline may already be too far underway for correction.

> For competition (exit) to work as a mechanism of recuperation from performance lapses, it is generally best for a firm to have a mixture of alert and inert customers. The alert customers provide the firm with a feedback mechanism which starts the effort at recuperation while the inert customers provide it with the time and dollar cushion needed for this effort to come to fruition.

Id. at 24. Several prominent class action scholars have discussed the problems of class action governance using Hirschman's work as a guide or inspiration. *See* Coffee, *Accountability, supra* note 3, at 376 n.17 (using the corporate governance paradigm of "exit, voice, loyalty" but asserting no intent to adopt specific conclusions from Hirschman's work); Samuel Issacharoff, *Governance and Legitimacy in the Law of Class Actions*, 1999 SUP. CT. REV. 337, 367 (1999).

61. The fairness hearing is non-adversarial where there are no objectors. Where there are objectors represented by sophisticated counsel, the objectors may bring an adversarial aspect into the proceeding.

court has not certified the class action.[62] Where a certified class is expanded during settlement negotiations, usually to assist defendants in obtaining global peace, the same problem arises.[63] Thus, even a settlement reached subsequent to certification may have a "settlement class"—a number of people whose first connection to the class action is through settlement and whose claims did not go through the certification process.[64]

Those due process protections that seem on their surface to encourage some class member participation have proven ineffective. The fundamental due process requirements for 23(b)(3) class actions are notice, the right to be heard, and the right to opt out.[65] Opt out, or exit, is an ineffective means of realizing participation values because small claims class members cannot bring independent claims. Notices are difficult for ordinary people to understand and hearings are non-adversarial and, therefore, fail to address inequities in settlements. The structure of notices and fairness hearings does not facilitate the possibility of persuasive objectors who are able to point out the inadequacy of settlements.

The means by which a class member may exit a class action is by opting out of the class either at the certification or settlement stage.[66] The central argument

---

62. *See* Issacharoff, *supra* note 60, at 348 (discussing court's treatment of settlement classes in *Amchem Products, Inc. v. Windsor*, 521 U.S. 591 (1997), and arguing that even though Supreme Court in *Amchem* warned against the evils of settling without reference to testing of claims in the adversarial process, that is just what settlement classes are). Because no certification motion has been approved in a settlement class, if defendants pull out of the settlement then the class will cease to exist. In that case, class counsel will have to litigate a difficult and costly certification motion.

63. *See* HENSLER ET AL., *supra* note 2, at 157-58 (discussing the expansion of the class after settlement in the Bausch & Lomb contact lens pricing litigation). That RAND study notes the expansion of class parameters in several of the case studies.

64. To the extent that we believe that the expansion of the class for settlement purposes does not raise notice problems, it is because we do not believe that the initial notice itself was particularly effective.

65. *See* Phillips Petroleum Co. v. Shutts, 472 U.S. 797, 812 (1985) (holding that for personal jurisdiction over absent class members, due process requires notice, the opportunity to be heard, and the opportunity to opt out); *see also* Patrick Woolley, *Rethinking the Adequacy of Adequate Representation*, 75 TEX. L. REV. 571 (1997) (arguing that due process requires notice, the right to be heard, and the right to opt out in the class action context and that these are inadequately met under the current regime); Monaghan, *supra* note 6, at 1166 (discussing due process requirements under *Shutts*).

66. *See* Fed. R. Civ. P. 23(c)(2)(B) (requiring "the best notice practicable under the circumstances, including individual notice to all members who can be identified through reasonable effort" and requiring that this notice be in "plain, easily understood language" and include "the nature of the action, the definition of the class certified, the class claims, issues, or defenses, that a class member may enter an appearance through counsel if the member so desires, that the court will exclude from the class any member who requests exclusion, stating when and how members may elect to be excluded, and the binding effect of the class judgment on class members under Rule 23(c)(3)"). Opt outs are not required in limited fund class actions brought under Rule 23(b)(1) or

in favor of opt outs is that they permit the realization of individual choice.[67] Courts understand opt outs as a pure expression of claimant autonomy, one that cannot be collectivized.[68] The autonomy rationale for opt outs assumes that claimants make knowing and intelligent decisions to opt out or stay in the class action, which we know is often not the case in small claims class actions. Individual opt outs make the most sense where a claimant does not want to be part of a lawsuit in the first place and/or has a positive value claim that can be brought independently, because opt outs place the claimant in the same position she would have been in absent the class action mechanism.[69]

Exit is not a viable means of realizing autonomy in a non-competitive space, however, because practically speaking, opting out simply removes that class member from the debate. Opt outs are not viable solutions to an inadequate settlement because once a class member has opted out, he or she no longer has standing to participate in the class action.[70] The opt out, as an expression of autonomy, could be valuable where there is a realistic alternative to the settlement; otherwise, the right to opt out is a mere formality.[71] A class member who has opted out of a small claims class action will not be able to file a claim on her own. From the perspective of that class member, the right to opt out is meaningless because there is nowhere to opt out *to*.

Theoretically, massive opt outs might be an effective means of communicating client dissatisfaction since courts consider not opting out to be

---

injunctive class actions brought under Rule 23(b)(2).

67. *See* Coffee, *Accountability*, *supra* note 3, at 418 (arguing that exit through opt outs maximizes individual choice).

68. "The right to participate, or to opt out, is an individual one and should not be made by the class representative or the class counsel." Hanlon v. Chrysler Corp., 150 F.3d 1011, 1024 (9th Cir. 1998). *Hanlon* argued that the right to opt out must be an individual right not only for due process concerns of the individual class members "who have the right to intelligently and individually choose whether to continue in a suit as class members," *id.*, but also because any other rule "'would lead to chaos in the management of class actions.'" *Id.* (quoting Berry Petroleum Co. v. Adams & Peck, 518 F.2d 402, 412 (2d Cir. 1975)). Collective opt outs, which have been rejected by courts, make more sense where the choice of law renders one group of claimants more valuable than others.

69. *But see* Rubenstein, *supra* note 14, at 1648-49 (arguing that the class device is inadequate because a litigant cannot get a case un-filed, and the opportunity to participate in remedy not useful for one who does not want the case at all). There is a possibility, albeit unlikely, that a class of opt outs could seek certification and mount a rival class action, and that the courts might harness such potential rivalries to yield more fair results. John Coffee proposes mechanisms for encouraging such competition. His solution is discussed *infra* pp. 32-33.

70. *See* Mayfield v. Barr, 985 F.2d 1090, 1093 (D.C. Cir. 1993) (finding that class members who have opted out of a damages class action have no standing to object to a subsequent class settlement because by opting out they "escape[] the binding effect of the class settlement").

71. *See, e.g.*, Coffee, *Accountability*, *supra* note 3, at 378 (discussing small claims class action where "right to exit will mean little" in contrast to the mass tort or large claims case where right is critical because at least some claimants hold high value claims).

tacit consent.[72] A near-total opt out would be something like a boycott of the settlement and would make possible a second, perhaps better, class action on behalf of opt outs. But in a world where numerous objections have not been sufficient to spur courts to reject settlements, a campaign of opt outs is not likely to be effective.[73] Courts may read massive opt outs as a protest, or may see them as an exercise in client autonomy that legitimates acceptance of the offending settlement for the remaining class members. This creates a double-bind for class members, who lack standing to intervene in the fairness hearing if they have opted out and are trapped in the settlement they disapprove of if they stay in, object, and lose. Even under the revised rule, which permits opt outs at settlement, the timing of the opt out is such that class members must choose whether to opt out before the fairness hearing.[74] Courts have rejected opt out attempts to appeal settlements for lack of standing, perhaps because courts sympathize with class counsel's arguments that attorneys representing opt outs are attempting to hijack the intervention process to obtain leverage over class counsel and defendants in order to procure a better settlement for the claimants who opted out.[75]

One may ask why we should be concerned about the exclusion of dissatisfied class members with small value claims. Some argue that we should not because absent the class action collectivization mechanism, these dissenters would have no recourse at all,[76] and therefore any recovery is a windfall for them. There are several responses to this view. First, because some dissatisfied class members

---

72. *See, e.g.*, Bell Atl. Corp. v. Bolger, 2 F.3d 1304, 1314 n.15 (3d Cir. 1993) (noting that "silence constitutes tacit consent," but recognizing that this assumption may, as a practical matter, understate potential objectors); *In re* GNC S'holder Litig., 668 F. Supp. 450, 451 (W.D. Pa. 1987) (noting that "[i]n the class action context, silence may be construed as assent").

73. *See, e.g.*, Reed v. Gen. Motors Corp., 703 F.2d 170, 174-75 (5th Cir. 1983) (approving settlement over objections of twenty-three of twenty-seven named plaintiffs and forty percent of class); TBK Partners, Ltd. v. W. Union Corp., 675 F.2d 456 (2d Cir. 1982) (approving settlement over objections of approximately fifty-six percent of class); Cotton v. Hinton, 559 F.2d 1326, 1333-34 (5th Cir. 1977) (approving settlement despite objections of counsel purporting to represent nearly fifty percent of class); Bryan v. Pittsburgh Plate Glass Co., 494 F.2d 799 (3d Cir.) (approving settlement despite objections of twenty percent of the class), *cert. denied*, 419 U.S. 900 (1974). *But see In re* Corrugated Container Antitrust Litig., 643 F.2d 195, 217-18 (5th Cir. 1981) (recognizing that "a low level of vociferous objection is not necessarily synonymous with jubilant support"); Pettway v. Am. Cast Iron Pipe Co., 576 F.2d 1157 (5th Cir. 1978) (overturning approval of settlement where seventy percent of class objected).

74. *See* FED. R. CIV. P. 23(e)(3) ("[T]he court may refuse to approve a settlement unless it affords a new opportunity to request exclusion to individual class members who had an earlier opportunity to request exclusion but did not do so.").

75. *See, e.g.*, *In re* Brand Name Prescription Drugs Antitrust Litig., 115 F.3d 456 (7th Cir. 1997) (barring organized opt outs from appealing settlement).

76. For example, Monaghan approves of the result in "situations where the absent plaintiffs lack independently viable claims and stand only to 'win' by the lawsuit." Monaghan, *supra* note 6, at 1169.

opt out for good reasons, the argument for ignoring opt outs is in effect an argument against the policy of vindicating and deterring small wrongs through class actions. The "windfall" argument also ignores the regulatory function of class action suits. Because legislatures have made consumer protection through lawsuits a staple of our political landscape, it is important that these suits produce fair results, whether that means individual compensation or deterrence. The deterrence value of small claims class actions further requires that the system ensure that the amount extracted from defendants is appropriate.[77] The deterrence value alone cannot determine distribution of the fund, however. There may be better justifications for excluding dissenters from decisions about how settlement funds are distributed, because they do not have a stake in that distribution.[78] On the other hand, dissenters will often be those who are most aware of the inequities of the distribution scheme developed by the attorneys, and their input will benefit class members who do stand to gain from the fund. In the context where the state has successfully consolidated claims through the class action mechanism in the hands of class counsel, dissenting voices become more important as indicators of a failure of process because competition is not a viable outlet for dissent.[79]

The two other procedural protections meant to create some limited class member participation, notice, and hearing, are also inadequate as currently enforced. As a form of communication with informational intermediaries and monitors, such as objectors' attorneys and judges, notices provide insufficient information. As a form of communication with class members, notices are difficult to understand and therefore ineffective. The types of notice used in most class actions indicate that courts rarely take the notice requirement seriously. Notices are generally written in a form of legalese that is difficult, if not impossible, for the ordinary claimant to understand.[80] Some are even

---

77. *See generally* Alexander, *supra* note 22 (arguing in favor of settlements that reflect outcomes at trial rather than the "going rate" of settlement in securities context).

78. This argument does not apply to claimants who stand to benefit from the fund. The only justification to exclude small claims class members from participating in such distribution decisions is that this type of input is too expensive or inconvenient to obtain.

79. Sometimes class members opt out because they oppose the action altogether as harmful to the interests of the class as a whole. In that case, opting out will not help them because they cannot get their case "un-filed" by opting out. *See* Rubenstein, *supra* note 14, at 1648-49 (noting that the class action device is an inadequate solution to the problem of the limits of individual autonomy in the group litigation context because class members cannot get a case "un-filed" and the opportunity to participate is not a useful remedy where plaintiffs do not want the lawsuit brought in the first place). *See, e.g.*, *In re* Matzo Food Prods. Litig., 156 F.R.D. 600, 608 (D. N.J. 1994) (rejecting settlement and noting that plaintiffs who have business relationships with defendant may be reticent to bring claim against defendant).

80. "It is beyond the experience or expectation of reasonable citizens that the failure to respond to what looks like a slightly unusual piece of junk mail constitutes assent to the solicitation . . ." Paul D. Carrington & Derek P. Apanovitch, *The Constitutional Limits of Judicial Rulemaking: The Illegitimacy of Mass-Tort Settlements Negotiated Under Federal Rule 23*, 39 ARIZ. L. REV.

inaccessible to a reader trained as an attorney.[81] Even the notice requirements of the revised Rule 23(c), which provide some specific disclosures, are insufficient.[82] Rule 23 does not require notices to include the amount an individual class member will receive or the amount of attorneys' fees.[83]

Fairness hearings are equally flawed. They are often too short to achieve the goal of genuine debate over the merits of the settlement.[84] The timing of filings further limits the usefulness of fairness hearings. Objectors are often required to file their opposition motions *before* class counsel and defendants file their motions in support of settlement. [85] This timing, combined with the limits on objector discovery, leaves objectors at a disadvantage because they must develop their objections without the information possessed by class counsel and defendants.[86] Such hearings provide little opportunity for meaningful objection and allow class counsel and defendants to push their settlement through without

---

461, 467-68 (1997).

81. *See, e.g.*, Kamilewics v. Bank of Boston Corp., 100 F.3d 1348, 1352 (7th Cir. 1996) (Easterbrook, J., dissenting) ("The notice not only didn't alert the absent class members to the impending loss but also pulled the wool over the state judge's eyes.")

82. *See* Fed. R. Civ. P. 23(c)(2)(B) (requiring that notices to class members in 23(b)(3) class actions provide the following information in "plain, easily understood language": (1) nature of the action, (2) the definition of the class certified, (3) the class claims, issues or defenses, (4) that a class member may enter an appearance through counsel, (5) that any class member may request exclusion from the class, (6) the binding effect of the class judgment).

83. *See* HENSLER ET AL., *supra* note 2, at 451-53. There is some movement towards more rigorous notice requirements. *See* State v. Homeside Lending, Inc., 826 A.2d 997 (Vt. 2003) (finding lack of personal jurisdiction over Vermont plaintiffs in part because notice was substantively inadequate); *infra* note 270 (citing cases holding that disclosure of attorneys fees is a required part of class notice).

84. *See, e.g.*, Wolfman & Morrison, *supra* note 5, at 489 (describing the Mustang convertible coupon settlement where objectors obtained no information prior to the fairness hearing and the judge approved settlement after a thirty minute hearing with no evidentiary support in favor of the settlement presented).

85. *See id.* at 480-90 (discussing the difficulties of obtaining information as objectors); *see also* Brief of Amicus Curiae Public Citizen in *In re* Prudential Ins. Co. of Am. Sales Practices Litig. Agent Actions, No. 99-5960 (3d Cir.), *available at* http://www.citizen.org/litigation/briefs/Class_Action/ articles.cfm?ID=693 (detailed discussion of legal and practical difficulties faced by settlement objectors).

86. Potential objectors are kept out of the negotiation process and are permitted very limited discovery. Courts have barred objectors from cross-examining fairness experts and from obtaining discovery of side-settlements. *See, e.g.*, Rutter & Wilbanks Corp. v. Shell Oil Co., 314 F.3d 1180, 1186 (10th Cir. 2002) (affirming settlement and holding that hearing did not violate due process where objectors were forbidden from cross examining parties' fairness expert, not permitted to present live testimony, and not permitted to present rebuttal affidavits); Duhaime v. John Hancock Mut. Life Ins. Co., 183 F.3d 1 (1st Cir. 1999) (denying objectors discovery of side-settlement). The revised Rule 23(e)(2) requires parties to file a statement identifying side deals but not the substance of those deals.

debate.

In their Rule 23(e) review of settlements for procedural fairness, courts will generally look to whether "(1) the negotiations occurred at arms length; (2) there was sufficient discovery; (3) the proponents of the settlement are experienced in similar litigation; and (4) only a small fraction of the class objected."[87] All but one of these criteria are fundamentally flawed. First, even parties engaging in arms-length negotiations can face incentives that lead them to self-serving settlements that do not adequately compensate the class. This is especially true where the "client" cannot monitor the attorney's actions. Second, in evaluating discovery, courts only look at discovery as between defendant's and plaintiff's counsel. Discovery is only sufficient if both class counsel *and* objectors have an adequate opportunity to obtain material information. Third, the absence of objectors has no necessary relationship to the fairness of a settlement. Few objectors can mean that there is little to object to because the settlement is fair. But it can also mean that class members lacked sufficient opportunity to object. For example, there may be few objections because notices were too confusing, class members had insufficient information, or class members were not given an adequate opportunity to voice their objections. In addition, objections may be limited because even though a settlement is unfair, class members have made the cost-benefit calculation that their potential individual recovery is too small to merit involvement. Other types of objectors, such as public interest groups or state attorneys general, may not object to unfair settlements because they did not know of the settlement or lacked the resources to intervene. Because the costs of objecting will exceed the value of any individual claim, unscrupulous class counsel can take advantage of the same type of collective action problem that was to be remedied by the class action device in the first place. Only the experience of the attorney in similar litigation (essentially a measure of quality of leadership), is a good procedural measure. Even this measure is problematic, however, because a self-dealing counsel may have substantial experience in similar litigation, but not have reached adequate results for her clients.

Compounding the flaws in these procedural requirements is the problem of ascertaining class member preferences. The structure of Rule 23 implies that class member preferences can be ascertained through the attorney-class relationship or, as a last resort, at a fairness hearing, but these provisions are inadequate to the task. There are numerous practical barriers to ascertaining class member preferences. Reaching out to class members takes attorney time and money. Voting, for example, would exponentially increase the costs associated with notice as each additional communication means substantial additional expense. Community meetings are also difficult to organize when class members are geographically disbursed. Even effective mechanisms might not guarantee high response rates.[88] Even if costs were not an issue, there are

---

87. *See In re* Gen. Motors Pick-Up Truck Fuel Tank Litig., 55 F.3d 768, 785 (1995).

88. There is little empirical data on claimant response rates, although commentators believe that they are low. *See* Rubenstein, *supra* note 14, at 1657-58 ("[F]ew class members respond to court mailings and those who do are not representative.").

other difficulties in ascertaining preferences. The standard economic assumption that class members' preferences are static and that claimants seek to maximize the payout on their claims is not entirely appropriate where claimants are not educated about their claims or where claimants are not engaged in repeated, similar transactions. Moreover, the settlements themselves are confusing and difficult to understand. Even if a decisionmaker was assured that claim maximization was the goal, substantive preference questions still arise because the possibilities for structuring payouts make for difficult choices. For example, reasonable claimants may differ as to whether a coupon or the promise of in-kind services or a *cy pres* fund is preferable to a cash payout of less value.[89]

The substance of preferences collected depends on the level of education provided to class members and the methods by which preferences are ascertained. For example, class members may wish to realize values other than claim maximization; they may prefer to support consumer advocacy groups rather than obtain minimal individual compensation.[90] If the formulator of the question assumes that claim maximization is the only possible value, however, class members will not be able to express this preference. Because class members' preferences will depend on how informed they are, the institutional designer needs to determine how much education to provide, balancing the cost of education against the value of better decision making by class members.[91]

Even minimal participation demands informed class members, but under current practices, as a result of incomprehensible notices, uninformative fairness hearings, and alienated representatives, class members are uninformed. Thus, even if class members were willing to monitor their agents, they could not do so effectively, and courts do not fill the breach. For example, neither the Rules nor most courts require the parties to report on the ultimate payout at the end of settlement administration.[92] They do not bar reversion of settlement funds to the

---

89. As one political scientist explained: "[u]nder the standard assumptions of modern economic theory, public choice theorists have demonstrated the logical impossibility of constructing *any* attractive, consistent procedure for making collective choices from unrestricted sets of three or more alternatives." LARRY M. BARTELS, PRESIDENTIAL PRIMARIES AND THE DYNAMICS OF PUBLIC CHOICE 296 (Princeton 1988). For an interesting discussion of nonpecuniary settlements and recommendations for evaluating such settlements, see Geoffrey P. Miller & Lori S. Singer, *Nonpecuniary Class Action Settlements*, 60 LAW & CONTEMP. PROBS. 97 (1997). The consumer class action settlements that have been most criticized tend to be coupon settlements in which class members receive the dubious opportunity to obtain more services from the defendant who cheated them in the first place. *See, e.g.*, Wolfman & Morrison, *supra* note 5, at 472-77 (discussing inadequate coupon settlements).

90. *See* Patricia Sturdevant, *Using the Cy Pres Doctrine to Fund Consumer Advocacy*, 33 TRIAL 80 (November 1997) (arguing in favor of distributing residue of class settlements to consumer advocacy organizations); Kerry Barnett, Note, *Equitable Trusts: An Effective Remedy in Consumer Class Actions*, 96 YALE L. J. 1591 (1987).

91. Bartels argues that students of governance should be concerned not with "summation of preferences" but with "formation of preferences." BARTELS, *supra* note 89, at 299.

92. For example, although the court in the *In re Mexico Money Transfer Litigation* retained

defendant to the extent that class members do not collect payment. Nor are "smooth sailing" agreements, in which defendants agree not to oppose class counsel's attorney's fee motion, forbidden.

Courts should ensure that an absence of objections is the result of well-educated class members being satisfied, rather than the result of objections being silenced.[93] This means encouraging objections, because class counsel cannot be relied upon to provide the court with necessary information to evaluate the settlement.[94] For example, in the *General Motors Pick-Up Truck Fuel Tank Litigation*, a settlement was nearly approved that would have provided class members with essentially non-transferable coupons for $1000 off of the purchase of a new GM truck.[95] These coupons disfavored the two groups most likely to own the offending truck, class members who lack the funds to buy a new truck,

---

jurisdiction over the settlement to review the injunction consented to by the defendants, the court did not order the parties to return to demonstrate that the settlement had in fact achieved the predicted results. 164 F. Supp. 2d 1002, 1033-34 (N.D. Ill. 2000). Such decisions can create an incentive to exaggerate the predicted settlement amount to increase class counsel's fee award where attorney's fees are calculated on a percentage of the fund. For example, in *Roberts v. Bausch & Lomb, Inc.*, No. CV-94-C-1144-W (N.D. Ala. filed Nov. 1, 1994) discussed in HENSLER ET AL., *supra* note 2, at 145-73, the total settlement was presented to the court as approximately $68 million depending on how many class members claimed rewards. The attorneys were paid $8 million based on that representation, calculating a fee of a little less than fifteen percent. The parties were not required to report the ultimate payout to the court, but based on SEC filings, RAND researchers deduced that the defendant never allocated more than $37.7 million to pay out all expenses relating to the litigation, including attorney's fees and administration costs. Thus there was likely a substantial overpayment to class counsel.

93. For example, in *Roberts v. Bausch & Lomb, Inc.*, there were no objectors. *Id.* Two other lawsuits had been filed in federal court in California and state court in New York, but rather than yielding objectors, the competing suits were settled for undisclosed amounts of attorneys' fees. *See* HENSLER ET AL., *supra* note 2, at 158-61. In that case, absent class members had no means of organizing an opposition or improvement to the settlement. This lack of participation may have had an effect on the settlement and on the award of attorneys' fees. The RAND study estimated that class members received a little over 30% of the fund. *Id.* at 429. *See also* Koniak & Cohen, *supra* note 9, at 1083-84 (describing BancBoston settlement with only one objector where class members were required to pay attorneys' fees in excess of their individual recoveries).

94. Class counsel and defendants' counsel cannot be relied on to point out injustice—they are equally likely to "put one over on the court." Kamilewicz v. Bank of Boston Corp., 100 F.3d 1348 (7th Cir. 1996) (Easterbrook, J., dissenting from denial of rehearing en banc), *quoted with approval in* Amchem Prods., Inc. v. Windsor, 117 S. Ct. 2231, 2249 (1997). *See In re* Ford Motor Co. Bronco II Prod. Liab. Litig., 1995 U.S. Dist. Lexis 3507, at *19 (E.D. La. Mar 15, 1995) (settlement rejected as having a value of "effectively zero" based on involvement by objectors); Bloyed v. Gen. Motors Corp., 881 S.W.2d 422 (Tex. App. 1994), *aff'd and remanded*, 916 S.W.2d 949 (Tex. 1996) (Texas Supreme Court's rejection of settlement as unfair based in large part on brief filed by objectors, Public Citizen Litigation Group and Center for Auto Safety).

95. *See In re* Gen. Motors Pick-Up Truck Fuel Tank Prod. Liab. Litig., 55 F.3d 768 (3d Cir. 1995).

and local governments that had fleets of trucks but were not likely to buy more during the applicable period.[96] The involvement of objectors and their appeal resulted in the settlement being overturned.[97] Thus, where responsible objectors have a serious and active role in the fairness hearing, they can often influence the result substantially.

Although in most cases the problem is too little objector involvement, objectors can also be the source of problems. Objectors' counsel have a perverse incentive to drop legitimate objections or soft-pedal them in order to obtain remuneration from the settlement because, to the extent that they are paid, objectors' counsel only receive payment if a settlement is approved.[98] As a result, objectors' counsel may try to use their leverage to extract undeserved rent from class counsel. Other times, objectors only marginally improve settlements but in a manner that also resembles self-dealing.[99] Fears of objectors multiplying proceedings or using their leverage to extract rent from class counsel have led some courts to impose significant sanctions against objectors.[100] As an empirical

---

96. *See* Wolfman & Morrison, *supra* note 5, at 472-77 (discussing the GM coupon settlement). Coupon settlements raise significant problems beyond the scope of this paper. Opponents of coupon settlements see coupon settlements as subject to serious abuse, especially when the coupons appear to be an advertisement for the defendant, and argue that consideration of whether a secondary market can be created in coupons should be required. *Id.* Other scholars believe that problems with coupon settlements can be overcome and that non-monetary settlements can be fair and efficient. *See* Miller & Singer, *supra* note 89. Miller testified in support of the *In re Mexico Money Transfer Litigation* coupon settlement on behalf of the defendants. 164 F. Supp. 2d at 1018-19 ("As Professor Miller explained to this court, a well designed coupon settlement can provide class members with more value than a cash settlement because the defendant is likely to be much more generous in its coupon offer.").

97. *See supra* note 95. The case was subsequently re-filed in Louisiana state court and the settlement affirmed with a revision that encouraged a secondary market in the coupons. *See* White v. Gen. Motors Corp., 835 So.2d 892 (La. Ct. App. 2002).

98. Most courts have held that objectors may obtain fees to the extent that their objections benefited the class. *See, e.g.*, Gottlieb v. Barry, 43 F.3d 474, 490-91 (10th Cir. 1994) (objector's fees appropriate where objectors "benefitted" the class).

99. For example, the objecting attorney in the *Mexico Money Transfer Litigation* was able to increase the ultimate settlement, but it is not clear how meaningful this increase was. This objecting attorney convinced the negotiating parties to make some substantive changes to the settlement, including adding a choice of one $6 coupon instead of only two coupons valued at $4.25, and doubling the *cy pres* donation from approximately two million dollars to four million dollars. *See In re* Mex. Money Transfer Litig., 164 F. Supp. 2d 1002, 1011 (N.D. Ill. 2000) (approving settlement that included *cy pres* fund where claimants were illegal immigrants thus difficult to find), discussed *supra* note 20.

100. *See* Vollmer v. Publishers Clearing House, 248 F.3d 698, 711 (7th Cir. 2001). In *Vollmer*, the district court levied $50,000 in sanctions—to be donated to a charity—against objecting attorneys on the basis of evidence that the objecting attorneys had no opportunity to dispute. Many of these sanctions are brought under 28 U.S.C. § 1927 (2003), which provides that

> [a]ny attorney or other person admitted to conduct cases in any court of the United

matter, these fears may be exaggerated and such sanctions may over-deter objector involvement or chill valuable objections.

D. The Limits of Judicial Policing

"[C]ourts have gone so far as to term the district judge in the settlement phase of a class action suit a fiduciary of the class, who is subject therefore to the high duty of care that the law requires of fiduciaries."[101] Nevertheless, as one prominent class action scholar has noted, "[p]erhaps in no other context do we find courts entering binding decrees with such a complete lack of access to quality information and so completely dependent on the parties who have the most to gain from favorable court action."[102] If judges approve a settlement in a case that they fundamentally believe is unsound,[103] or approve a poor settlement in a case they believe is strong,[104] then we may rightly be concerned about whether the judge is acting as an independent decisionmaker or granting judicial

---

> States or any Territory thereof who so multiplies the proceedings in any case unreasonably and vexatiously may be required by the court to satisfy personally the excess costs, expenses, and attorneys' fees reasonably incurred because of such conduct.

Id. See also In re Prudential Ins. Co. Am. Sales Practice Litig. Agent Actions, 278 F.3d 175, 193 (3d Cir. 2002) (upholding $100,000 in sanctions against objecting counsel and reversing on due process grounds a "scarlet letter" sanction requiring that objecting counsel attach documentation of the sanctions proceeding to all future *pro hac vice* applications). The Court upheld the sanctions against the objecting counsel despite the lack of express finding of bad faith below. *See id.* 194 (Rosenn, J., dissenting).

101. Reynolds v. Beneficial Nat'l Bank, 288 F.3d 277, 280 (7th Cir. 2002) (Posner, J.) (overturning settlement of class action on grounds that district court did not scrutinize sufficiently whether the settlement was collusive and failed to quantify the net expected value of continued litigation). *See also In re* Warner Communications Sec. Litig., 798 F.2d 35, 37 (2d Cir. 1986) ("a district court has the fiduciary responsibility of ensuring that the settlement is fair"); *In re* Cendant Corp. Litig., 264 F.3d 286, 296 (3d Cir. 2001) (same).

102. Issacharoff, *supra* note 6, at 808.

103. *See, e.g., In re* Mex. Money Transfer Litig., 267 F.3d 743, 748 (7th Cir. 2001) ("Were the class's claims worth more than $40 million, plus *cy pres* relief, plus the value of the injunction? Like the district court, we think not—indeed, we think that the claims had only nuisance value . . . .").

104. *Compare In re* Rhone-Poulenc Rorer Inc., 51 F.3d 1293 (7th Cir.), *cert. denied*, Grady v. Rhone-Poulenc Rorer Inc., 516 U.S. 867 (1995) (denying certification of blood products class as unmanageable), *with In re* Factor VIII or IX Concentrate Blood Prods. Litig., 159 F.3d 1016, 1020 (7th Cir. 1998), *cert. denied*, Mull & Mull, PLC v. Rhone-Poulenc Rorer, Inc., 526 U.S. 1081 (1999) (affirming approval of class-wide settlement of blood products claims and labeling settlement of individual cases with different levels of damages "downright weird"). This case is thoroughly described in HENSLER ET AL., *supra* note 2, at 293-317. The contrast between Judge Posner's opinion in the *Blood Products Litigation*, and Judge Posner's more recent decision in *Reynolds v. Beneficial National Bank*, 288 F.3d 277, 280-81 (7th Cir. 2002), in which he asserted that judges have a fiduciary duty towards class members, is stark.

imprimatur to an inadequate result, perhaps in violation of her fiduciary duty to the class.

The judge's role in class actions ranges from the traditional responsibility to evaluate and determine cases, to the more modern responsibilities to manage the interaction of various actors, monitor agents, determine the extent of plaintiff participation, and encourage the case's resolution.[105] Checks on district judges making fairness determinations are limited by the deferential standard of review employed by the appellate courts.[106] Although judges are required to accept or reject settlements as they are, there is the possibility for judges to reject settlements and provide guidance to the parties as to what a fair settlement might look like.[107] The very things that make a judge a more experienced arbiter of the fairness of a class action settlement—her experience in the difficulty of crafting it—may influence her to approve the settlement in order to complete a long and arduous litigation and/or negotiation.[108] If a judge is inclined to involve herself and assist in the crafting of a fair settlement, this may create a bias on the part of the judge in favor of the settlement that she worked to create.

Under the current regime, judicial accountability for the settlement results in an unreliable governance structure. The extent to which judges police class actions is largely a matter of the individual judge's choice, influenced by traditional views of the attorney-client relationship and settlement as a form of private ordering. Not only will judges not do enough to review settlements, there is little incentive for continued judicial oversight in the administration phase. There is no requirement in the Federal Rules for judicial review of settlement administration. Although judges retain jurisdiction over settlements of class actions until administration is complete, whether the settlement administration will be reviewed or a final report issued is at the discretion of the individual judge or party request.[109] This lack of uniformity is compounded by incentives for judges to approve settlements and not to maintain active oversight over their subsequent administration. Judges understandably may choose to focus their

---

105. One example of this is Judge Weinstein's case resolution methods, described in PETER SCHUCK, AGENT ORANGE ON TRIAL: MASS TOXIC DISASTERS IN THE COURTS (1987).

106. *See* Joel A. v. Giuliani, 218 F.3d 132, 139 (2d Cir. 2000) (standard of review of fairness determination is abuse of discretion).

107. "[I]t is not a district judge's job to dictate the terms of a class settlement; he should approve or disapprove a prosed agreement as it is placed before him and should not take it upon himself to modify its terms." *In re* Warner Communication Secs. Litig., 798 F.2d 35, 37 (2d Cir. 1986). *See also* Hanlon v. Chrysler Corp., 150 F.3d 1011, 1026 (9th Cir. 1998) ("The settlement must stand or fall in its entirety.").

108. *See, e.g.*, Twelve John Does v. District of Columbia, 117 F.3d 571, 576 (D.C. Cir. 1997) (approving a settlement under the abuse of discretion standard and finding that "the district court's experience overseeing the case for nearly two decades" gave it "a unique familiarity with the issues and the performance of class counsel").

109. *See* Tyson v. City of New York, 97 Civ. 3762 (JSM), Submission of Proposed Stipulation of Settlement (Jan. 5, 2000) (on file with author) (requiring that administration of claims be completed before attorneys fees are awarded and providing for interim fees).

limited resources on cases currently being adjudicated. The prospect of continuing oversight when the parties have agreed to a settlement may strike some judges as superfluous.

Accordingly, judges are both a source of hope for class action governance, in that they are a ready-made monitor of self-interested class counsel, and a source of frustration, in that the protections afforded class members by judicial oversight are dependent on the discretion of the individual judge.

II. PROPOSED MODELS OF CLASS ACTION GOVERNANCE: THE LAST TWENTY YEARS

In the past twenty years, scholars have proposed various mechanisms for solving the problems of class action governance, particularly in the settlement phase where agent-principal problems are most acute and where judicial discretion plays a significant role. The proposed solutions to the various problems posed by class actions fall into three general categories: market solutions,[110] democracy-based solutions,[111] and judicial administrative solutions.[112] Each of these offers significant contributions to aspects of class action problems, but each is fundamentally incomplete.

A. Market Solutions

Scholars have proposed numerous market mechanisms to solve the problems posed by small claims class actions. In looking for an analogy to class actions, some scholars have turned to corporate law.[113] As in the corporate context, this theory maintains, the problem with class actions is an agent-principal problem,[114]

---

110. *See, e.g.*, Coffee, *Accountability*, *supra* note 3, at 370 (advocating additional opt out mechanisms, i.e., "exit").

111. *See, e.g.*, Deborah L. Rhode, *Class Conflicts in Class Actions*, 34 STAN. L. REV. 1183, 1185 (1982) (proposing "a theory of representation mandating full disclosure of, although not necessarily deference to, class sentiment").

112. *See, e.g.*, Wolfman & Morrison, *supra* note 5, at 439 (proposing amendments to Rule 23 that increase judicial scrutiny of settlements and create more substantive guidelines for evaluating settlements); Shapiro, *supra* note 21, at 917, 922-23 (advocating an "entity" approach to class actions).

113. *See, e.g.*, Charles Silver & Lynn Baker, *I Cut, You Choose: The Role of Plaintiffs' Counsel in Allocating Settlement Proceeds*, 84 VA. L. REV. 1465 (1998) (making an explicit analogy to the corporate context).

114. *See* Resnik et al., *supra* note 24, at 374 (describing how economic analysis of litigation has been based on the agent-principal problem); Coffee, *Accountability*, *supra* note 3, at 375 (describing class action as an "organizational form" involving a "principal/agent relationship"). *But see* Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24, at 683-84.

> Yet for analytical purposes, one better understands the behavior of the plaintiff's attorney in class and derivative actions if one views him not as an agent, but more as an entrepreneur who regards a litigation as a risky asset that requires continuing investment decisions. Furthermore, a purely fiduciary perspective is misleading because it assumes

arising from the fact that the class members' ownership of their claims is separated from control over those claims. Accordingly, scholars urge that class action reform requires policy-makers to "consider market-based remedies and checking mechanisms that have worked in related contexts to align the interests of the principal and the agent."[115] The two most prominent market mechanisms proposed in the class action literature concern the structure of attorney fee awards and the availability of opt outs.[116]

1. Attorneys' Fees.—Monitoring class counsel is central to class action governance. Legal economists have focused their attention on the incentive structure for class counsel to "pursue their own interests at the expense of the class" resulting in cheap settlements.[117] As courts and commentators have noted, class counsel have an incentive to engage in self-dealing regardless of their ethical obligations and duties to the class.[118] Some scholars propose creating

---

> that the client's preferences with respect to when an action should be settled are exogenously determined, when, in fact, they are largely influenced by the fee award formula adopted by the court.

115. Coffee, *Accountability*, *supra* note 3, at 371. Under this view, class members may be deemed to have consented to representation by the attorney where "the agency costs associated with the relationship have been minimized." *Id.* at 376. There have been numerous propositions for methods of aligning class counsel's interests with those of the class. *See, e.g.*, Christopher R. Leslie, *A Market-Based Approach to Coupon Settlements in Antitrust and Consumer Class Action Litigation*, 49 UCLA L. REV. 991, 1088 (2002) (advocating that class counsel be paid in the same currency as the class).

116. Some scholars have even proposed to do away with the class members altogether by auctioning off their claims to the highest bidder. *See* Macey & Miller, *supra* note 7, at 6 (proposing a class action regime whereby claims are auctioned off to attorneys). This proposal has generated an interesting scholarly discussion. *See, e.g.*, Randall S. Thomas & Robert G. Hansen, *Auctioning Class Action and Derivative Lawsuits: A Critical Analysis*, 87 NW. U. L. REV. 423, 424-25 (1993). Thomas & Hansen only address auctions of the right to *represent* the class, not auctions for purchase of class claims.

117. Staton v. Boeing Co., 313 F.3d 447, 467 n.12 (9th Cir. 2002), *withdrawn and reh'g denied*, 327 F.3d 938 (9th Cir. 2003).

> Even when there is no direct proof of explicit collusion, there is always the possibility in class action settlements that the defendant, class counsel, and class representatives will all pursue their own interests at the expense of the class. For that reason, the absence of direct proof of collusion does not reduce the need for careful review of the fairness of the settlement, particularly those aspects of the settlement that could constitute inducements to the participants in the negotiation to forego pursuit of class interests.

Id. See also Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24, at 669.

118. Courts have held that class counsel owes a fiduciary duty to the class as a whole. *See In re* Gen. Motors Corp. Pick-Up Truck Fuel Tank Prods. Liab. Litig., 55 F.3d 768, 801 (3d Cir. 1995) ("Beyond their ethical obligations to their clients, class attorneys, purporting to represent a class, also owe the entire class a fiduciary duty once the class complaint is filed."); Stewart v. Gen. Motors Corp., 756 F.2d 1285, 1294 n.5 (7th Cir. 1985) (citation omitted) ("[B]oth the class

competition among attorneys for the privilege of representing a class.[119] Others propose giving a powerful plaintiff the ability to negotiate a market rate on behalf of the class.[120] Still others propose fee regimes that reward class counsel incrementally, based on the actual benefits conferred to the class,[121] or attempt to solve the problem by subsidizing rival class actions amongst which class members can choose.[122]

Of the proposals put forward, the one most consistent with accepted attorney compensation norms is to link attorneys' fees to the amount of benefit the attorney provides the class, with a marginally increased percentage of the fund going to the attorney to encourage the maximization of claimholder value.[123] Large fee awards are beneficial because they create an incentive for attorneys to bring small claims suits as private attorneys general, and it makes sense to link that incentive directly to the benefits conferred to the class.[124] Yet, it also makes sense to limit percentage fee awards so that they bear some relationship to the effort expended by the attorneys and do not become a windfall to attorneys at the expense of the class. The problem is that there is little empirical evidence to help judges determine the appropriate percentage of the fund to give to attorneys to attract responsible and experienced attorneys to bring class action suits and to compensate for the risk that they incur in taking on such suits, without offending sensibilities. Setting a loadstar cap on attorneys' fees would mitigate the concern that attorneys receive a windfall, but even setting such a cap is difficult without more information on how the balance should be struck.

Another solution that has been much discussed in the literature is auctioning

---

determination and designation of counsel as class representative come through judicial determinations, and the attorney so benefited serves in something of a position of public trust.").

119. *See* Macey & Miller, *supra* note 7, at 6 (proposing auctions for the position of class counsel); Koniak & Cohen, *supra* note 9, at 1113-14 (advocating same and arguing that the absence of auctions may constitute an antitrust violation).

120. Jill E. Fisch, *Lawyers on the Auction Block: Evaluating the Selection of Class Counsel By Auction*, 102 COLUM. L. REV. 650 (2002) (criticizing auctions and proposing stronger lead plaintiffs as an alternative solution to the agent-principal problem).

121. *See* Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24, at 725. *Compare Third Circuit Task Force Report on Selection of Class Counsel*, 74 TEMP. L. REV. 689, 704-05 (2001) (advising that "[t]he traditional methods of selecting class counsel, with significant reliance on private ordering, are preferable to auctions in most class action cases. In using those traditional methods, however, the court must guard against overstaffing by lawyer groups."), *with* John C. Coffee, Jr., *Litigation Governance: A Gentle Critique of the Third Circuit Task Force Report*, 74 TEMP. L. REV. 805, 808 (2001) [hereinafter Coffee, *Litigation Governance*] (arguing that "private ordering works least well when agency costs are high and competition is limited." Also noting that "as a general rule, professional groups have little incentive to seek increased competition or lower fees, and professional ethical norms generally are enforced only against outliers and insurgents.").

122. *See* Coffee, *Accountability*, *supra* note 3, at 424.

123. *See* Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24, at 725.

124. *See id.* at 721-25.

off the right to represent a class.[125] In essence, the auction proposal entails potential class counsel submitting proposals for representation and fees at the beginning of the suit, and the court awarding the case to the lowest bidding counsel.[126] This proposal may reduce attorneys' fees, but will do little to eliminate agency costs. Attorneys may "settle early in order to obtain a larger profit on the fee"[127] or, if they are obtaining a percentage of the total fund, exaggerate the value of the settlement to increase their compensation. An auction may produce lower attorneys' fees, but may not obtain the best quality or most experienced counsel.[128] As their proponents concede, they cannot eliminate agency costs associated with the attorney-class relationship.[129] Thus, on their own terms, attorneys' fee mechanisms are an incomplete solution to class action problems.

The agent-principal analysis of class actions has led to an overemphasis on mechanisms to control attorneys' fees. An attorneys' fee solution to class action governance is analogous to a corporate governance scheme based solely on controlling executive compensation. The observation that various forms of self-dealing, including "cheap settlements," are the likely result of a flawed fee regime is accurate, but it captures only one aspect of the weak governance structure of class actions.

2. Opt Outs.—Opt outs are the only exit mechanism available to class members. Recognizing that the absence of competition diminishes the significance of opt outs, John Coffee has proposed the creation of a market in opt

---

125. *See* Macey & Miller, *supra* note 7, at 105-18 (describing proposals for auctioning off claims and for auctioning rights to represent the class).

126. *See, e.g., In re* Oracle Secs. Litig., 131 F.R.D. 688, 689-90 (N.D. Cal. 1990) (ruling that competitive bidding would determine selection of class counsel and counsel's compensation). *See also* Koniak & Cohen, *supra* note 9, at 1202-06 (arguing that market for representation of class should be competitive and that current system may be an antitrust violation). The revised Rule 23(g) offers something in this regard by permitting the court to choose class counsel. In a twist on this idea, Geoffrey Miller has also proposed auctions for class counsel after settlement has been reached, permitting rival class counsel to post a bond for the settlement amount and attempt to negotiate a better deal for the class. *See* Geoffrey P. Miller, *Competing Bids in Class Action Settlements*, 31 HOFSTRA L. REV. 633 (2003).

127. Macey & Miller, *supra* note 7, at 113.

128. *Id.* at 113; Lucian Arye Bebchuk, *The Questionable Case for Using Auctions to Select Lead Counsel*, 80 WASH. U. L.Q. 889, 892 (2002); Coffee, *Litigation Governance*, *supra* note 121. *See also* Fisch, *supra* note 120, at 652 (arguing that auctions are poor tools for selecting firms based on multiple criteria, that auctions compromise the judicial role, and that they are unlikely to produce reasonable fee awards as an empirical matter).

129. It does not seem rational that an attorney who has purchased the rights to a class action at a slightly lower fee will take the nuances of class interests into account any more than the attorney entitled to potentially exorbitant fees. One might even think that the auction attorney is less likely to take class desires seriously, so long as they do not affect the ability of the class action settlement to be approved or otherwise interfere with the total recovery, because he has an even greater sense of ownership over the lawsuit than the "joint venturing" class action attorney did.

outs rather like a hostile takeover in the corporate context.[130] When a settlement has been proposed, attorneys who wish to represent dissenting class members would be permitted to include a "counter-solicitation" in the notice of settlement sent to class members, inviting them to opt out of the settlement.[131] If successful, these competing attorneys would organize dissatisfied class members and file a competing class action.[132] According to Coffee, this approach would solve the abandonment problem of opt outs without threatening "the finality of settlements nor encourag[ing] collateral attacks on the class settlement."[133] It would be relatively cost efficient because the "counter-solicitation" could be included with the regular settlement notice.[134] Finally, it would provide greater opportunity for the substantive exercise of class member autonomy at least to the extent that the choices presented are reasonably attractive to the class members.

To create a viable market in opt outs, the potential gain from competition would have to be greater than the cost of the new attorney and the rival solicitation. We might predict that competition would be limited where a settlement is inadequate but the value that the rival attorney is able to extract from the sub-class is not high enough to merit the investment in organizing a rival solicitation. Without empirical analysis, it is difficult to tell whether that result is optimal in terms of the regulatory function of small claims class actions and whether the cost of mounting a rival action would be too high for attorney entrepreneurs.[135] It is worth noting that in the securities context, when sophisticated class members are dissatisfied with class counsel, instead of creating rival class actions by organizing opt outs, these potential lead plaintiffs prefer to opt out and file independent actions in state court.[136]

Coffee recognizes some of the problems with this creative proposal,[137] yet

---

130. *See* Coffee, *Accountability*, *supra* note 3, at 422.

131. *Id.* at 423.

132. *Id.*

133. *Id.* at 424.

134. *Id.* at 423-24.

135. Even if rival attorneys are permitted to solicit opt outs in settlement notices, the cost and risks of pursuing such a rival class action may be too high for attorney entrepreneurs in small claims actions. Interview with Joseph Grundfest, W.A. Franke Professor of Law and Business, Stanford Law School (Feb. 13, 2003).

136. *See, e.g.*, Cal. State Teachers' Ret. Sys. v. Qwest Communications Int'l, No. 415546 (Cal. Super. Ct., San Francisco County filed Dec. 10, 2002), *available at* http://www.sftc.org/ (state lawsuit with no class or federal claims filed to avoid class litigation); *UC Sues Firms with WorldCom Ties*, L.A. TIMES, Feb. 14, 2003, at C4 (Regents of the University of California, which had purchased 10.2 million shares of Worldcom between 1998 and 2000, pulled out of nationwide federal class action and commenced suit on their own.). These plaintiffs bring only state court claims to avoid removal and class adjudication.

137. Coffee notes that counter-solicitations create the danger that the class will become too fragmented as a result of significant number of proxy proposals, and proposes that the problem of too many proxy proposals might be solved by requiring insurgent counsel to pay reasonable costs of printing and mailing proposals (as the SEC handles hostile proxy solicitations). Counter-

leaves many questions unanswered. How would class members intelligently choose among the multi-factored proposals likely to arise and the associated risks? More than likely, such rival solicitations would require opt outs to give up the proverbial bird in the hand for a bird in the bush.[138] Would there be disclosure requirements that enabled class members to make these choices and, if so, how much disclosure or education would be necessary and in what form?[139]

Second, this market model would suffer the same agent-principal problems observed under the current regime. The incentives of class counsel and rival attorneys would continue to be the same: to maximize their own payout by settling early or for inadequate amounts. Rival attorneys may be bought out or co-opted, as happens in competing class actions under the current regime. Thus, even if there were such competition, agent monitoring would still be necessary. Moreover, this competitive model would further hamper the ability of the court to act as such a monitor and to evaluate the fairness of settlements because the process of soliciting opt outs would siphon objectors away to the new action rather than encourage them to voice their objections to the proposed settlement. In addition, judges would have a difficult time determining whether in any particular case a lack of competition for opt outs is the result of market inefficiencies or an indication that the settlement is indeed fair.[140]

A rising trend that bears some relationship to the rival solicitation proposal

---

solicitation could also be made contingent on rival counsel getting enough support, and the court could refuse to permit rival solicitation if too few class members sign on or otherwise fail tests of Rule 23. *See* Coffee, *Accountability*, *supra* note 3, at 426. Coffee also suggests that a court could arbitrarily limit the number of competing solicitations (to two or three) "leaving others to contact class members on their own if they wished." *Id.* at 426. On the other hand, if the number of rival solicitations are limited to avoid fragmentation, then the class may be denied the full panoply of choices that a true market would offer. *Id.* Rival attorneys may experience judicial hostility to a proposal intended to delay and complicate an existing settlement process. *Id.* at 427. Additionally, there may be substantial reputational costs for prestigious class action firms to mount counter-solicitations. In a world where class action attorneys are repeat players, it is not difficult to imagine that there may be costs meted upon dissident counsel. As a result, the type of experienced attorneys one would ideally want to run a rival class action may not be willing to create such a rivalry unless the gain is substantial enough to outweigh the potential reputational costs.

138. This differentiates the proposal from the hostile takeover context where the shareholders are offered specific sums of money for their shares: "the class member is still very much subject to risk, while the shareholder who accepts a cash tender offer is no longer subject to the risks of the enterprise." Coffee, *Accountability*, *supra* note 3, at 423 n.140.

139. Such a rival solicitation may also face other hurdles, such as when class dissatisfaction is based on non-monetary factors, including class counsel's relationship with the class, a desire for greater voice, or the desire for non-monetary relief out of which it would be difficult to take a contingency share.

140. Similarly, in the current regime it is difficult for courts to determine whether a small number of objectors is the result of the fact that there are too many barriers to objection or of the settlement being genuinely fair.

is the devolution of class actions to state courts.[141] Charles M. Tiebout's familiar model of local government provides a useful analogy.[142] Tiebout theorized that the distribution of public goods in a given community reflects local preferences because a consumer-voter will pick the community "which best satisfies his preference pattern for public goods."[143] The natural corollary of this proposition is that "[t]he greater the number of communities and the greater the variance among them, the closer the consumer will come to fully realizing his preference position."[144] If compensation through a class action is a public good, then it would seem that permitting class members a choice would maximize their ability to obtain a preferential benefit from the class action mechanism. But if the choice is a stark one between obtaining nothing (which is the normal result of opting out in most small claims class actions) and joining an inadequate settlement, the "choice" is only in name.

Applying Tiebout's theory to class actions, we can envision a regime that would encourage class actions on a local scale instead of collecting claims into nationwide class actions.[145] Local class actions would be more manageable and actions based on state laws would not suffer from choice of law problems. Localization would make public monitoring by the states' attorneys general easier. It would also create a kind of competitive environment, allowing courts to compare the terms of settlements with those of similar class actions in other states. On the other hand, this type of "market" in class actions, demarcated along state boundaries, does not offer class members greater choice and autonomy. Moreover, localization would reduce the efficiencies achieved by national class actions. Finally, further localization may not be feasible at a time when policy makers, attorneys and judges, are increasingly enlarging class actions, expanding defendants' ability to remove such actions to federal court and settle them as nationwide class actions rather than permitting them to be split off by jurisdiction.[146]

The central limitation of exit mechanisms as a sole solution to unsatisfactory settlements is the same problem as all exit mechanisms in markets with limited competition share: limited and ineffective opportunities to opt out benefit the

---

141. *See supra* note 44 and accompanying text (describing state class actions limited to citizens of forum state).

142. *See* Charles M. Tiebout, *A Pure Theory of Local Expenditures*, 64 J. POL. ECON. 416 (1956). *See, e.g.*, William W. Bratton & Joseph A. McCahery, *The New Economics of Jurisdictional Competition: Devolutionary Federalism in a Second-Best World*, 86 GEO. L. J. 201 (1997) (criticizing Tiebout's model); Gerald E. Frug, *City Services*, 73 N.Y.U. L. REV. 23, 25-33 (1998) (criticizing Tiebout's model among other things for the unrealistic assumption that some people choose poor service communities and noting limits of mobility in the real world).

143. Tiebout, *supra* note 142, at 418.

144. *Id.*

145. *See supra* note 44 (discussing nascent trend towards localized class actions).

146. One example is the Class Action Fairness Act of 2003, S. 274, 108th Cong. § 4 (2003), discussed *supra* note 45. It seems clear that legislation advocating unchecked removal of class actions to federal courts is not necessarily protective of class members, but rather favors defendants.

decisionmakers, who are spared the trouble of addressing genuine concerns because their most vociferous opponents have left the stage.[147] It is more beneficial for the architects of a class-wide settlement to permit some opt outs as a kind of safety valve to get rid of dissenting voices, than to allow a realistic option that might scuttle settlement. Thus we often see courts validating settlements on the basis that class members had an opportunity to opt out and chose not to do so, without looking to the role of the opt out as a mechanism for minimizing dissent.[148] Like attorneys' fee regimes, exit mechanisms alone are not a satisfactory solution to the governance problems in class actions.

These types of market mechanisms ignore a central difference between agent principal problems in the class action context as opposed to the corporate context. For claimants the class action is a single, unique transaction, and subsequent market penalties for singular instances of mismanagement are inadequate.[149] The complex problems in class action governance cannot be resolved by resort to a single incentive mechanism because class actions are not merely a function of the plaintiffs' attorney. A simple market model would not work even if there were vigorous competition between class counsel bringing identical class actions—which in itself creates obvious inefficiencies—because the panacea of attorney competition overlooks the role of other actors in class action governance, including judges, class members, objectors, and opt-outs. These observations represent a significant hurdle to any market solution to class action problems.[150]

B. Democratic Solutions

Another source for class action governance mechanisms is democratic theory. Rule 23(b)(3) provides a number of mechanisms that resemble political

---

147. "Those who hold power in the lazy monopoly may actually have an interest in *creating* some limited opportunities for exit on the part of those whose voice might be uncomfortable." HIRSCHMAN, *supra* note 60, at 60. Some limited competition may actually comfort monopolist by taking away the most vocal customers. *Id.* at 59.

148. *See* discussion *supra* at note 73 (discussing court treatment of opt outs and objections in settlement context); *see also* Devlin v. Scardelletti, 536 U.S. 1, 122 S. Ct. 2005, 2011 (2002) (noting that because class member was denied opt out, appeal gained greater importance for him to protect himself).

149. *See* FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 103 (1991) (discussing the ineffectiveness of market mechanisms on singular breaches of fiduciary duty in the corporate context). In the corporate context, voting and purchasing mechanisms, as well as boards of directors, discipline management. Shares are fungible, and regulation attempts to create transparency of information on which shareholders base their investments. That is simply not the case in non-fungible small claims class actions where claimants do not "opt in"; there are no disclosure mechanisms and only the judge, who is not elected by the claimants, disciplines class counsel.

150. This is not to say that market solutions cannot cure some of what ails class actions, but that market mechanisms alone do not provide a complete governance system.

devices. The class as a group has a representative in the proceedings.[151] The fairness hearing provides a forum where the merits of settlement proposals are presented and discussed prior to the judge's decision. In deciding whether to approve a settlement, judges take into account certain displays of popular will from the class, such as numerous opt outs or objections to the settlement.[152]

Scholars have proposed democratic governance structures for class actions in the context of civil rights injunctive actions.[153] However, deliberative solutions have not been a prominent part of the discourse concerning consumer class actions. This may be because consumer class actions are viewed more as an economic matter than a political one, whereas civil rights actions relate directly to the political realm. Viewed in this light, it may seem hypocritical to permit experts to run a civil rights action in order to vindicate democratic values, particularly when the underlying claim relates to some form of disenfranchisement from the political or social sphere. Perhaps deliberation in the civil rights context is seen to "rectify the antidemocratic exclusion of chronically disadvantaged groups from the theatre of politics."[154]

Democratic values have relevance, however, even to complaints based on economic harms. The observation that "[d]emocratic decisionmaking is an attractive alternative to unrestrained liberty because it provides a means for reining in the self-appointed community representative; it also checks the alienation and disempowerment that result from overreliance on lawyers"[155] applies just as well to the consumer class action context. The mandatory nature of injunctive civil rights actions is not a sufficient reason to limit democratic proposals to them because opt outs do not provide a true alternative in the small claims context. Furthermore, there is nothing inherent in consumer issues, as opposed to civil rights issues, that render them less amenable to class member input into their resolution. Although consumer class actions generally address small individual claims, they are nevertheless significant protective and preventive mechanisms in a world where interactions with economically and politically powerful corporations are such a pervasive aspect of people's daily

---

151. *See* FED. R. CIV. P. 23(a)(4). Although it is not clear what "representation" means in this context, as discussed below, this provision indicates that some importance is attached to the voice of class members.

152. *See supra* text accompanying note 65-70 (discussing the judicial approach to opt outs) and note 72-73 (discussing the judicial approach to objectors).

153. Examples of discussions of the role of democratic decision making processes in civil rights class actions include Rhode, *supra* note 111, at 1191-92 and Rubenstein, *supra* note 14, at 1663-68. My research has not revealed substantial scholarship making a consistent argument for a democratic approach of any kind to class actions outside the area of civil rights injunctive class actions. *Cf.* Elizabeth J. Cabraser, *Enforcing the Social Compact Through Representative Litigation*, 33 CONN. L. REV. 1239 (2001) (discussing the representative litigation as an aspect of the realization of the American social contract).

154. Rubenstein, *supra* note 14, at 1659 (quoting Jane S. Schacter, *Metademocracy: The Changing Structure of Legitimacy in Statutory Interpretation*, 108 HARV. L. REV. 593, 619 (1995)).

155. *Id.* at 1659-60.

life.

In the mass tort context, scholars' discussions of participation values have been limited to support or rejection of the idea of universal, particularized adjudication.[156] However, the argument that small claims (or even mass tort) plaintiffs do not have a process right to individual participation that trumps the right of the public or defendant to conserve resources,[157] as one prominent scholar has argued, should not necessarily mean that policy makers should ignore methods for *collective* input from the class into the settlement process. The question is whether it is possible to realize democratic values in the small claims class action in the absence of individualized adjudication.

It may be inherently good for collective settlements to reflect democratic processes because they are political; their regulatory role makes them so.[158] Furthermore, settlements can be preference forming as well as preference expressive vehicles. They can create expectations as to the class' specific claims and future suits.[159] Just because the class action settlement is a transaction does not mean that it needs to be viewed as a purely private transaction with no political dimension or as limited to class counsel and defendants.

1. Representation.—The Rule 23(a)(4) requirement of "adequate representation" raises an obvious comparison to representative democracy. Although self-appointed representatives are more in keeping with dictatorial systems of government than with representative democracy,[160] as political theorists remind us, representation "need have nothing to do with popular self-government."[161] A representative may merely be someone with the authority to act on behalf of her constituency.[162]

Scholars and courts have struggled to create a thicker description of what

---

156. *See* Rosenberg, *supra* note 19, at 211-12 (arguing against the idea that self-determination requires individualized litigation); *see also* Judith Resnik, *From "Cases" to "Litigation,"* 54 LAW & CONTEMP. PROBS. 5, 9-14 (1991) (discussing the 1966 Advisory Committee's concerns about applying class action to mass torts).

157. *See* Rosenberg, *supra* note 19, at 213 n.6.

158. Frank Michelman, *Law's Republic*, 97 YALE L. J. 1493, 1503 (1988) (discussing political participation as a positive human good).

159. *See* Alexander, *supra* note 22, at 500 (criticizing settlements in securities class actions that are settled at a "going rate" rather than based on the strength of the parties' claims).

160. *See* Fiss, *supra* note 25, at 25.

161. HANNA F. PITKIN, THE CONCEPT OF REPRESENTATION 2 (1972).

162. To the extent one can argue that class representatives represent at all, it is in the Burkean conception of the representative as a trustee without accountability. As Burke wrote, "[t]he king is the representative of the people; so are the lords; so are the judges. They are all trustees for the people." Quoted in PITKIN, *supra* note 161, at 129. Under the current damages class action regime, the representative appears to lack both the authority to act, because the lawyer is in fact the representative with authority, and has no accountability to the class membership. It is not clear, under this scenario, where the content of the notion of "class representative" is to be derived. A class representative chosen by an attorney or self-appointed, without the power to make decisions, is a "representative" only in the thinnest sense.

role representatives take in the class context and of the source, if any, of their legitimacy.[163] Some scholars have defined the legitimacy of the representative in the class context as inhering in her ability to be a "mirror" of the class.[164] She is a representative in the sense that she has suffered the same injury or asserts the same claims as absent class members. This view stands behind the logic of the cohesion theory of representation.[165] Cohesion theory assumes that if the claims are identical, the class representative will necessarily follow the intent of the class as a whole.[166]

This conception is unsatisfactory for several reasons. First, it assumes that the representative is imbued with the power to control the litigation. Under our

---

163. *See* Rhode, *supra* note 111, at 1191-92. Rhode describes how, because of class certification, the requirement that representative parties will fairly and adequately represent the class is of constitutional dimension, but that this requirement does not tell litigants what it means to "adequately protect" the "interests" of the class and asks, "Do the named representative and counsel serve primarily as 'instructed delegates' pursuing objectives to which a majority of class members subscribe?" *Id.* at 1192. Or, in the alternative, "does the representative role track Edmund Burke's notion of an 'enlightened trustee,' who makes an independent assessment of class concerns?" *Id.* at 1192-93. She concludes that courts have "done little more than acknowledge the absence of any 'clear principles governing the allocation of decisionmaking authority between the attorney and the class.'" *Id.* at 1193. Twenty years later, this appears to still be the case. *See also* Coffee, *Accountability*, *supra* note 3, at 373 (critiquing the Supreme Court's treatment of the notion of representativeness and class cohesion as overly formalistic). The Burkean notion of the representative as a trustee with special expertise (that is, the lawyer) is often rejected as overly paternalistic and not in keeping with ideas of litigation as an autonomous act. For a sustained discussion of the expertise model in this context, see Rubenstein, *supra* note 14, at 1163-68. For a thoughtful historical discussion of the relationship between representation and participation, see Robert G. Bone, *Personal and Impersonal Litigative Forms: Reconceiving the History of Adjudicative Representation*, 70 B.U. L. REV. 213, 297-98 (1990) (reviewing STEPHEN C. YEAZELL, FROM MEDIEVAL GROUP LITIGATION TO THE MODERN CLASS ACTION) (concluding that 23(b)(3) class actions in some ways resemble participation-based representative suits).

164. *See* Coffee, *Accountability*, *supra* note 3, at 375 (discussing the representative as "mirror image").

165. *See* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 621 (1997) (whether class should be divided into multiple, independently represented sub-classes depends on whether the class "has sufficient unity so that absent members can fairly be bound by decisions of class representatives"). This conception originates in the requirements of commonality and typicality found in Rule 23(a). Although on its facts *Amchem* is a mass torts case, its language provides an overarching ruling concerning the basis of representation in class actions generally.

166. It is not clear how proponents of the cohesion principle believe that the translation from the individual representatives interests to group interests occurs—whether the representative's decisions will reflect the decisions of each individual member or will reflect the range of decisions that the groups many members would have made. What is clear is that the cohesion principle is based on the theory that the right of representation is "not a day in court but the right to have one's interest adequately represented." Owen M. Fiss, *The Allure of Individualism*, 78 IOWA L. REV. 965, 970-71 (1993).

current regime, the real authority to make decisions on behalf of the class rests with the class' attorney.[167] Significant additional oversight mechanisms would have to be instituted to make the class representative an effective monitor of attorney conduct. Second, it requires a very high level of identity between the representative and the class. Such a level of identity may not be entirely possible, nor entirely necessary to good governance. It is difficult to draw the line between those individual qualities of class members that require identity and those qualities that do not matter for purposes of representation. As a corollary, the Rule 23 typicality requirement, together with the representation requirement, places reliance on the court's ability to determine that the class representative is in fact like the class in all material ways and will make rational choices on their behalf. This determination is, to put it mildly, an inexact science and no provision is made for situations in which the representatives' differences from other members of the class subsequently become material.[168] Finally, the cohesion principle attempts to define away the problem of inter-class discord by removing from the discourse the individuals who may disagree.

It is difficult to imagine a class achieving this strong coherence requirement and still retaining the efficiency of the class action device.[169] Cohesion and efficiency are in tension with one another. The looser the concept of cohesion, the more efficiently claims can be disposed of through the class action mechanism. The tighter the concept of cohesion, the more difficult it is to forge a class action. There are two equally unpalatable ways to achieve cohesion: remove individuals with different interests from the class, thereby narrowing it to a near impossibility, or redefine their interests in a cribbed and unrealistic way. Either way, it does not provide an underlying rational for the legitimacy of any one particular group identity or for the particular representative appointed to defend its interests.

2. Deliberation.—Some scholars have turned to democratic deliberation as a model for listening to class members. In exploring the idea of deliberation in the class context, Deborah Rhode presented two models: pluralism and majoritarianism.[170] The pluralism model consisted of individualized representation for discrete constituencies within the class.[171] As commentators have noted, successive sub-classification has the potential to fragment the class

---

167. *See supra* note 6 (describing the view that class counsel is the true architect of class actions).

168. For example, what if the representative is lazy, risk averse or a risk seeker? *See* Coffee, *Accountability*, *supra* note 3, at 375 (discussing potential problems with representative).

169. *See* discussion *supra*, accompanying notes 17-25 (discussing the efficiency basis for class action litigation).

170. *See* Rhode, *supra* note 111, at 1185.

171. *See id.* This is similar to the model that was eventually adopted by the Supreme Court in mass tort cases under the guise of adherence to the cohesion principle. *See* Ortiz v. Fibreboard Corp., 527 U.S. 815, 856 (1999) (requiring sub-classification where sub groups within class have conflicts of interests).

action out of existence.[172] On the other hand, this model also has the potential to create small, more manageable and responsive, class actions. Without more development as to what governance of sub-classes means, this proposal suffers from the same problems of the cohesion principle; it assumes that a smaller scale will cure the multiple problems of class actions.

The majoritarian model would ascertain members' preferences through polls or plebiscites.[173] But within this model, some entity must still shape and ask the questions and oversee the referendum process.[174] Majoritarianism also assumes that preferences are clear and ascertainable, not contentious, disparate and dynamic, and just need to be "collected," although this conflicts with most people's experience.[175] Furthermore, there are practical barriers to plebiscites, such as the natural apathy of class members with small stakes in the litigation[176] and the cost of voting mechanisms as well as voter education.[177] Commentators

---

172. *See* Coffee, *Accountability*, *supra* note 3, at 374.

173. *See* Rhode, *supra* note 111, at 1185.

174. Not every aspect of a case is amenable to voting or deliberation, such as purely strategic legal decisions that would be made by the attorney alone in a traditional attorney-client relationship. *See* Rubenstein, *supra* note 14, at 1654 ("Democratic values would lend significant legitimacy to goal-based decisions, but would have less applicability to the more technical decisions about legal strategies."). Furthermore, as Rhode herself points out, public choice models demonstrate how agenda setting can shape results. *See* Rhode, *supra* note 111, at 1236. *See also* Rubenstein, *supra* note 14, at 1656-57 (discussing hurdles to voting). In addition, there is what Rhode called the "unarticulated premise" against majority vote, that "[t]he class as an entity has interests beyond those expressed by its current constituents." Rhode, *supra* note 111, at 1241. In consumer class actions, third parties may have an interest in the outcome of the class action. For example, the public has an interest in the adequacy of deterrence and in avoiding over-deterrence.

175. *See* discussion supra note 88-89 (describing the difficulty of ascertaining class member preferences).

176. One of the most important and largely unresolved hurdles to a plebiscite is participation. Scholars have recognized the difficulties of getting class members to respond to communications from class counsel, especially where the stakes are relatively small. "Few class members respond to court mailings and those who do are not representative." Rubenstein, *supra* note 14, at 1657-58. *See also* Paul D. Carrington & Derek P. Apanovitch, *The Constitutional Limits of Judicial Rulemaking: The Legitimacy of Mass-Tort Settlements Negotiated Under Federal Rule 23*, 39 ARIZ. L. REV. 461, 466 n.35 (1997) ("The likelihood that a class member will actually receive and comprehend the notice of the action is in every case very small. Frequently, the cost of reading and understanding the notice exceeds the benefits, and not infrequently, the notice is impenetrable by the average citizen."). These scholars have had difficulty resolving the lack of participation without resorting to a model based on attorney expertise. Rubenstein, *supra* note 14, at 1663-68. Rubenstein argues that the expertise model has been given short shrift, and that attorney expertise has its place in group litigation, such as in making procedural or technical decisions. *Id.*

177. One cannot easily disaggregate the difficulty of obtaining class member participation from the types of participation class members have been permitted. Attorneys have incentives to keep class members out of the negotiation and settlement process and notices are often mind-numbingly difficult to read. These factors must be added to the cost-benefit analysis of class member

have also expressed the paternalistic concern that class members will not know what is best for them, or will choose the wrong result out of shortsightedness.[178]

Another approach is consultation-based deliberation.[179] One concrete proposal is a procedural 'community consultation' requirement for class actions.[180] Such a procedural requirement may help to alleviate concerns about coercion in the beginning of the class action lawsuit, although it would require procedural reform and would likely be difficult to implement.[181] By what criteria will decisionmakers determine the sufficiency of consultation? This type of solution is particularly difficult to implement in consumer class actions because of the lack of a coherent "community" with which to consult.

There are some analogies to community consultation in the context of small claims class actions. Courts might seek out consumer advocacy groups or, if one demographic is particularly affected, "community representatives" of that group, although that concept is problematic as well. This has occasionally occurred in the context of *cy pres* settlements. In *In re Mexico Money Transfer Litigation*, for example, a case concerning a class made up almost entirely of Mexican-Americans and Mexican residents of the United States who transferred funds to their families in Mexico, the district court heard testimony from a number of representatives of Mexican communities in the affected areas concerning the four million dollar *cy pres* fund.[182] In that case, both the objectors and class counsel brought forward community representatives, the end result of which was

---

participation.

178. *See* Rhode, *supra* note 111, at 1237 (articulating concerns that class members' positions may be due to posturing or reflect peer pressure, or that class members views may not result in the best outcome because they are too risk averse or risk seeking).

179. In seeking an alternative to exclusive attorney decision-making in the civil rights context, Rhode advocated a "theory of representation mandating full disclosure of, although not necessarily deference to, class sentiment." *Id.* at 1185.

180. Rubenstein seeks to use procedural and ethical rules to "promote more democratic means of client goal-setting and more expertise-driven norms of attorney decisionmaking in group litigation." Rubenstein, *supra* note 14, at 1668. He proposes rules requiring lawyers to show that they had some community interaction prior to their decision to file, or to show community participation or that democratically elected representatives approve of the action. *Id.* at 1659. For example, he suggests a kind of "community derivative suit," in which the representative would have to file a special pleading showing that the case "grew out of some prefiling democratic processes" and if the action did not "flow from a democratically produced group decision" it could be dismissed. *Id.* at 1670-71. Rubenstein recognizes that this deliberative model, which offers only limited participation, would be difficult to implement and would potentially create an overwhelming barrier to suits.

181. Rubenstein notes the problems with the "community derivative" idea are that there is no literal "community" with a principal place of business, the idea of community likened to corporations could lead to less democracy because some sub-groups might get undue power and the procedural hurdle would be costly and might undermine civil rights suits generally. *See id.* at 1672-73.

182. 164 F. Supp. 2d 1002, 1020-21 (N.D. Ill. 2000), *aff'd*, 267 F.3d 743 (7th Cir. 2001).

something similar to a battle of the experts.[183] Although such community involvement can lead to consensus, it may also reflect or even increase differences within the group. Such consultation can deteriorate into a battle of competing community leaders, taking the role of experts rather than authentic voices of the community. Nevertheless, disputes should not always be construed as negative. The inability to create consensus around the settlement may not be the product of failed judicial leadership, but a reflection of political reality.[184]

C. The Judicial Administration Model

Another group of scholars and judges subscribe to a model of class action governance based on judicial monitoring.[185] Many of these proposals have been inspired by mass tort cases.[186] The subtext of most proposals based on judicial monitoring is that the wrongs vindicated through the class action mechanism are best addressed through governmental regulation, but in the absence of legislative action, the class action requires courts to act like administrative agencies.[187] The central feature of these proposals, strengthening the hands of the court, is not limited to the mass tort context.

1. The Entity Model.—One group of scholars has described an "entity theory" of class actions where the class is understood as an entity rather than an aggregate of individuals.[188] Entity theory is not a specific governance mechanism, but rather a reconceptualization of class structure that would change the way procedural rules are applied in class actions. It proposes to abandon autonomy-based approaches to class actions and shift the focus of the Rule 23(a) inquiry on the class as a whole. Scholars have suggested that the entity theory would improve the current law on class actions by rationalizing the mootness

---

183. *Id.*

184. In *Hart v. Community School Board*, 383 F. Supp. 699 (E.D.N.Y. 1974), Judge Weinstein did not order the special masters' community redevelopment plan but did exhort third parties to follow the plan. *See id.* at 775; Curtis J. Berger, *Away from the Court House and into the Field: The Odyssey of a Special Master*, 78 COLUM. L. REV. 707 (1978) (describing his experience as a special master in the *Hart* case).

185. *See, e.g.*, Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 630-41 (1997) (Breyer, J., dissenting); JACK B. WEINSTEIN, INDIVIDUAL JUSTICE IN MASS TORT LITIGATION: THE EFFECT OF CLASS ACTIONS, CONSOLIDATIONS AND MULTIPARTY DEVICES (1995); Edward H. Cooper, *Rule 23: Challenges to the Rulemaking Process*, 71 N.Y.U. L. REV. 13 (1996); Shapiro, *supra* note 21, at 920.

186. *See, e.g.*, Shapiro, *supra* note 21, at 920 (focusing on mass tort because "this has been the principal area of current debate").

187. *See Amchem Prods., Inc.*, 521 U.S. at 599 ("In the face of legislative inaction, the federal courts—lacking authority to replace state tort systems with a national toxic tort compensation regime—endeavored to work with the procedural tools available to improve management of federal asbestos litigation.").

188. The entity theory seems to have been first proposed by Edward Cooper. *See* Cooper, *supra* note 185, at 13.

doctrine, focusing on the actual claim rather than the individual representative, and encouraging a more direct focus on the class and potential conflicts requiring sub-classification.[189]

Entity theory would strengthen the trend towards federalization and nationalization of class actions. Under the entity theory, the amount in controversy rules would apply to the class as a whole rather than the individual litigants, and courts could determine whether the amount in controversy has been met on an aggregate basis.[190] The entity theory might also affect the way we think about due process in the class context, making "membership in the litigating class itself a tie to the forum."[191] An entity approach may also dictate that a single law should apply to all class members, thus eliminating the choice of law issues.[192] Although it seems difficult to imagine that the procedural device of the class action should affect the law governing an individual located in his home state, judges approve settlements that do not account for differences in the laws of individual states.[193] Most importantly, according to Cooper, an entity approach to Rule 23 would focus the court's inquiry more acutely on the adequacy of the attorney's representation, where it rightfully belongs.[194] Thus, the entity theory follows up on scholarship that observed that the attorney entrepreneur is the true engine of class action litigation.[195]

The entity theory was further developed by David Shapiro, who argued that "the wisest and most efficient way of promoting individual justice" is through collective mechanisms such as the class action lawsuit.[196] The viability of the entity theory in a given context depends on the cohesion of the class and the extent to which the class' interests are coextensive.[197] This justification for collective treatment explicitly abandons procedural autonomy values in favor of a rough-justice approach to individual claims.[198] It also depends on the purposes

---

189. *Id.* at 28.

190. *See id.* This is in contrast to an approach to the amount in controversy requirement permitting jurisdiction where it is met by the named plaintiff, like the rule for diversity jurisdiction in class actions. *See* Supreme Tribe of Ben Hur v. Cauble, 255 U.S. 356 (1921); Class Action Fairness Act of 2003, S. 274, 108th Cong. § 4 (2003), discussed *supra* note 45.

191. *See* Cooper, *supra* note 185, at 29.

192. *See id.* (describing a uniform choice of law mandate as too extreme and stating, "[a]s a mere procedural device, class treatment cannot alter the conceptual substance of the individual claim, no matter how drastically the claim is affected in fact").

193. *See, e.g.*, Wolfman & Morrison, *supra* note 5, at 460-61 (discussing approval of settlements despite substantial choice of law issues).

194. *See* Cooper, *supra* note 185, at 31.

195. *See* Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24.

196. *See* Shapiro, *supra* note 21, at 916.

197. *See id.* at 922-23.

198. From a political theory perspective, this view might be described as communitarian because greater weight is given to the common good than to the realization of individual rights, based on the idea that justice for individuals is only realized through common good. *See generally* MICHAEL SANDEL, DEMOCRACY'S DISCONTENT: AMERICA IN SEARCH OF A PUBLIC PHILOSOPHY

of the substantive law at issue. Where deterrence is the primary goal, Shapiro argues, an entity approach is more palatable because compensation inherently entails an individualist orientation.[199]

The entity theory would "severely limit[] such aspects of individual autonomy as the range of choice to move in or out of the class or to be represented before the court by counsel entirely of one's own selection."[200] Shapiro proposes that the class might be represented by a group of members who would receive notice and the opportunity to object, in place of the individualized notice currently required.[201] The focus of the district court's procedural due process inquiry would shift radically towards the adequacy of the representative and away from notice and opt out requirements.[202] Shapiro justifies this call for limited notice and opt outs with a very narrow view of due process rights in the class action context.[203] But he recognizes that this proposal would require, at a minimum, a revision of Rule 23.[204] Curtailing the notice requirement would prevent the expense of notice from becoming a barrier for small claims that "deserve[] recognition as a matter of substantive law" and "individual stakes seem to be worth the cost of litigating" but are too small to merit the notice

---

(1996).

199. *See* Shapiro, *supra* note 21, at 923-31 (arguing that the main purpose of small claims class actions is deterrence, and therefore the entity approach makes sense).

200. *Id.* at 919.

201. Shapiro argues that the Supreme Court's interpretation of Rule 23 requires more robust notice, but that due process concerns may be more flexible. *See id.* at 937 (citing Mullane v. Cent. Hanover Bank & Trust Co., 339 U.S. 306 (1950)). This requires a very strained reading of *Phillips Petroleum Co. v. Shutts*, 472 U.S. 797 (1985). *See* Monaghan, *supra* note 6, at 1159-63. *See generally* Patrick Woolley, *Rethinking the Adequacy of Adequate Representation*, 75 TEX. L. REV. 571, 573 (1997) (arguing that individual notice, hearing and the opportunity to opt out are due process requirements).

202. *See* Shapiro, *supra* note 21, at 937.

203. Shapiro's view is that because the Supreme Court did not focus on opt outs and notice in *Supreme Tribe of Ben Hur v. Cauble*, 255 U.S. 356 (1921), and *Hansberry v. Lee*, 311 U.S. 32 (1940), but instead on adequacy of representation as a due process threshold, the former are not truly due process requirements. *See* Shapiro, *supra* note 21, at 938, 958-59. Shapiro dismisses *Shutts* as essentially dictum and argues that the notice and opt out requirements articulated in that case may not be constitutionally required. *See* Shapiro, *supra* note 21, at 938, 954-55.

204. Although he argues that due process does not prohibit a looser view of notice and opt out, Shapiro believes that using Rule 23 to move towards an entity model would overstep the Court's bounds under the Rules Enabling Act and the Rules Decision Act. *See* Shapiro, *supra* note 21, at 953. In this regard Shapiro focuses on the requirements that individual notice be sent out upon class certification and opt out rights, and argues that the Rule

> should be framed in a way that does not place unreasonable roadblocks in the way of movement toward an entity model by responsible policymakers, nor should it impede recognition of the present force and effect of the model in the administration of class actions. At present, the rule may well fail both of these criteria.

Id. at 957.

costs.[205]

Shapiro advocates a revision of Rule 23 to create stronger judicial oversight of the adequacy of counsel, focusing on the experience of counsel, potential conflicts of interests, and the availability of channels of communication between counsel and a representative group of the class members. [206] Judicial review at the 23(e) hearing would then focus on the fairness of settlements overall and as to individual class members. It is not clear from Shapiro's vague description how this revision would be more than a codification of the current regime without opt out or notice protections.[207] In comparison to the entities recognized in our legal system, Shapiro's governance solution seems too simple.[208] For example, corporations or trade unions have governance mechanisms to select and monitor directors and to regulate membership participation. Corporations, which share some of the agent-principal problems of class actions, are heavily regulated, and subject to mandatory disclosure requirements, proxy voting and market forces.[209]

From an implementation perspective, the entity theory faces significant hurdles, particularly because notice and opt out requirements are well entrenched in due process jurisprudence and in Rule 23. There are, however, some doctrinal developments that seem to move closer to the entity model. These include the recent revision of Rule 23 to require judicial approval of class counsel,[210] an increasing focus on the class cohesion by the Supreme Court,[211] increased attention to the conflicts of interests of groups within the class,[212] and the development of a doctrine of judicial fiduciary duty to the class as a reaction to conflicts of interests problems with class counsel.[213]

---

205. *Id.* at 956.

206. *See id.* at 959. This approach holds promise but is insufficiently developed.

207. Shapiro supports his view with reference to other corporate bodies, such as a trade union that may be authorized by the majority of workers to represent them in bargaining or litigation even when an individual worker may not have agreed to that representation. *See id.* at 921. The weakness of the entity theory as a solution to the problem of class action governance is well illustrated by the example of a trade union. While majority vote is an acceptable democratic means of picking a representative, self-selection is not. As Shapiro himself points this out in another example, that of a municipality, "a company town organized and run by one's employer has less to be said in its favor than a truly public municipality." *Id.* at 922 n.18. While Shapiro focuses on this problem in the context of a defendant-defined class, it is equally a problem in a class defined by class counsel.

208. *See* John Leubsdorf, *Pluralizing the Client-Lawyer Relationship*, 77 CORNELL L. REV. 825, 828 (1992) (criticizing entity theories of attorney-client relations that do not provide decision-making arrangements for groups).

209. Recent experience with the failure of corporate governance seems to indicate that more, not less, watchfulness is necessary in the corporate sphere.

210. *See* proposed FED. R. CIV. P. 23(g), described *supra* note 6.

211. *See* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 621 (1997).

212. *Id.* at 626.

213. *See* Reynolds v. Beneficial Nat'l Bank, 288 F.3d 277, 279-80 (7th Cir. 2002) (overturning settlement of class action on grounds that district court did not scrutinize sufficiently whether the

Despite its problems, the entity theory has its appeal. Thinking of the class as a group, rather than as individual litigation multiplied, brings us a step closer to developing a coherent theory of class action governance. But the entity theory needs to do more than reinforce the dictatorial system currently in place.[214] Its focus on the client, a reference to the traditional system of individual litigation, is also a call for a coherent system of governance that will serve the same control and monitoring functions in the class context that the client does in the individual litigation context. As currently described, however, entity theory does not fulfill its promise.

2. Relying on Judges: Justice Breyer and Judge Weinstein.—While no judge has advocated adherence to the entity theory of class actions described above, there are examples of judicial support for a more collectivist or even communitarian view of the class action. Judicial focus in this regard is more on fair outcomes than fair process of group litigation. Two prominent judges have in their opinions and writings adopted an administrative view of group litigation that departs from the traditional atomistic approach: Justice Breyer and Judge Weinstein. Although these judges have considered the administrative model in the mass tort context, their thinking has implications for consumer class actions.

Justice Breyer's concurrence in part and dissent in part in *Amchem Products, Inc. v. Windsor*,[215] argued that, despite the flaws the majority found in the structure of that settlement, because the tort system was not working and because atomized litigation caused "[d]elays, high costs and a random pattern of noncompensation," a judicially created system of compensation through settlement was appropriate for asbestos claims.[216] He essentially advocated the creation of a temporary administrative agency with limited appellate review.[217] Accordingly, he was prepared to review settlements on an abuse of discretion standard that would be significantly more deferential to the district court's decisions than the majority.[218] Justice Breyer's view acknowledges the extent to

---

settlement was collusive and failed to quantify the net expected value of continued litigation and stated that "[w]e and other courts have gone so far as to term the district judge in the settlement phase of a class action suit a fiduciary of the class, who is subject therefore to the high duty of care that the law requires of fiduciaries.").

214. As Cooper notes, it is hard to pretend that the class is a "real" client. "The need for a client is most real in cases that aggregate large numbers of small claims and do not win the involvement of any class members with substantial stakes." Cooper, *supra* note 185, at 31-32.

215. 521 U.S. 591 (1997).

216. *Id.* at 632 (Breyer, J., dissenting in part).

217. *See generally* Martha Minow, *Judge for the Situation: Judge Jack Weinstein, Creator of Temporary Administrative Agencies*, 97 COLUM. L. REV. 2010 (1997) (proposing that Judge Weinstein's work on mass tort settlements essentially creates temporary administrative agencies).

218. *See* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 631 (1997). In reaching this decision, Justice Breyer underscored the special nature of this settlement as a compensation system preserving money for future illnesses and characterizes the settlement as "unusual in terms of its importance, both to many potential plaintiffs and to defendants, and with respect to the time, effort and expenditure that it reflects." *Id.* at 633. But despite this mitigating language, he appears to

which settlement is a political act, worth the sacrifice of individualist autonomy values for the good of the group.

This argument favors almost any settlement because the alternative can always be painted as very limited or long delayed compensation. Justice Breyer advocated judging the adequacy of representation by results; if the settlement is not patently unfair the higher courts should defer to the lower court's view on the settlement.[219] By focusing on the fairness of the settlement terms overall, courts shift their attention away from individuals or subgroup conflicts within the class. But this emphasis on outcomes raises serious questions about process. In the class action device, who is the decisionmaker most competent to make fairness determinations? Do we believe that we can evaluate the fairness of a result without looking to the fairness of the process by which it was reached? In the alternative, is an unsatisfactory outcome reached by full and fair process acceptable?

By contrast, the majority opinion in *Amchem* focused on individual interests.[220] How should this differential between individual rights and group compensation be reconciled? Because this tension will always have to be traded off to some extent in class action settlements, perhaps courts should look at the degree to which individual interests have been sacrificed in favor of those of the collective. Because the push towards settlement may be especially strong in difficult and complex cases, we should be concerned when settlement approval depends on a district judge's idiosyncratic estimation of fairness.[221]

---

advocate the deferential standard of review for *all* settlements. It is not difficult to argue that every class action presents unique, intractably difficult problems requiring solution at the expense of individual fairness. *Cf.* Richard A. Nagareda, *Turning from Tort to Administration*, 94 MICH. L. REV. 899, 902-03 (1996) (urging the use of the "hard look" doctrine in cases where appellate courts review class action settlements).

219. Justice Breyer describes himself as "agnostic" about the basic fairness of the *Amchem* settlement. 521 U.S. at 639.

220. *See id.* at 629 (Rule 23, "applied with the interests of absent class members in close view," cannot carry the load of this type of administrative model.). Although *Amchem* was a mass torts, positive claim class action, the Court did not distinguish its reasoning from other types of class actions, such as the small claims class actions discussed here.

221. As Justice Ginsburg wrote for the majority in *Amchem*, "the standards set for the protection of absent class members serve to inhibit appraisals of the chancellor's foot kind—class certifications dependent upon the court's gestalt judgment or overarching impression of the settlements' fairness." *Id.* at 621. While in his dissent Justice Breyer accused the Court of looking too deeply into the fairness of the settlement, it appears that he himself had also made a determination as to its fairness in light of the state of asbestos litigation. The factors courts consider in deciding whether to approve settlements, set forth in *City of Detroit v. Grinnell Corp.*, 495 F.2d 448 (2d Cir. 1974) and widely used by circuit courts across the country, set forth concrete standards of evaluating settlement fairness. Nevertheless, as Judge Posner's affirmance of the approval of the Blood Factor settlement indicates, these factors do not cabin judges as much as perhaps they should. *See In re* Factor VII or IX Concentrate Blood Prods. Litig., 159 F.3d 1016, 1018 (7th Cir. 1998) (affirming settlement but describing it as "downright weird").

Within the administrative model there is also a countervailing focus on process and contractarianism. Justice Breyer's insistence on the deferential standard of review bespeaks a trust in district court judges, perhaps subscribing to views expressed by some circuits that judges have a fiduciary duty towards the class.[222] To the extent that there is a divergence between the interests of the judicial system and the interests of the class, however, judicial fiduciary duty may not be enough. Justice Breyer, for example, expresses trust in the process of negotiation between the two sides and focuses on the work that plaintiffs' attorneys put into the settlement in that case.[223] Unfortunately, the difficulty of negotiations does not guarantee an absence of conflict between plaintiffs' attorneys and the class, nor eliminate the structural incentives for self-dealing.[224]

To be effective, judicial administration requires a free flow of accurate and complete information to the judge to exercise her expertise. In fact, however, there are often significant limitations on the information provided to the judge.[225] There is a danger, therefore, that the judge will only rely on the parties for information and be denied any opposing views of facts and law. Although some judges view their role as that of a fiduciary of the class, others view settlement as a type of private ordering over which their oversight role is limited. This variation creates unpredictability, a concern from a governance perspective.

Because it leaves control in the hands of a single actor, the judicial administration model raises some of the same concerns as other expertise models, that it is an "elitist subversion of democratic equality," infringing on individual liberty and simultaneously limiting the self-realization of the class as a group.[226]

---

222. *Amchem*, 521 U.S. at 640.

223. *Id.* at 633.

224. The *Amchem* settlement structure raised very real concerns that class counsel might sell out one portion of the class—or the class as a whole—in favor of their individual clients or, worse yet, in favor of current money and current fees in exchange for small future awards. There was also the specter that attorneys could cut off whole classes of claims in the interest of creating global settlement or simply to settle. The unique problems presented by the settlement addressed in *Amchem* are discussed in Susan P. Koniak, *Feasting While the Widow Weeps:* Georgine v. Amchem Products, Inc., 80 CORNELL L. REV. 1045 (1995), and Wolfman & Morrison, *supra* note 5, at 449-59. For a discussion on the unique aspects of the asbestos litigation, see Deborah R. Hensler, *As Time Goes By: Asbestos Litigation After* Amchem *and* Ortiz, 80 TEX. L. REV. 1899 (2002).

225. "Perhaps in no other context do we find courts entering binding decrees with such a complete lack of access to quality information and so completely dependent on the parties who have the most to gain from favorable court action." Issacharoff, *supra* note 7, at 808. This is compounded by the failure of some courts to permit objector discovery and their expressed desire not to delve too deeply into settlement. *See, e.g.*, Mars Steel Corp. v. Continental Ill. Nat'l Bank & Trust Co., 834 F.2d 677, 684 (7th Cir. 1987) ("The temptation to convert a settlement hearing into a full trial on the merits must be resisted.") (citing Airline Stewards & Stewardesses Ass'n v. Am. Airlines, Inc., 573 F.2d 960, 963-64 (7th Cir. 1978)). We should be concerned that case law suggests that discovery rules ought not apply to settlement hearings because the rules "eliminate the efficiency gained by the settlement itself." Wolfman & Morrison, *supra* note 5, at 485.

226. Rubenstein, *supra* note 14, at 1664.

It may result in the loss of vigorous discourse because power is concentrated in the hands of the judge.[227] Moreover, some have argued that this model raises separation of powers issues to the extent that the judge usurps the power of the legislature or executive to create what is essentially a temporary claims administration body.[228]

On the other hand, a judicial expertise model has its benefits. It centralizes and unifies decision-making, may lead to better quality and more neutral decisions, and ensure correct allocation of resources.[229] This model also solves the problems of atomization and the difficulty of creating mechanisms for participation by giving power to a neutral decisionmaker without a monetary stake in the outcome of the proceedings.[230] Still, some practical questions remain. What kind of duties does such a judge have towards the class? What kind of flexibility of standards and transfer of power is required to enable the judge to truly live up to this important role? And would the judge's increased involvement in settlements compromise judicial independence? These and other questions have been raised concerning the work and thought of Judge Jack Weinstein, who is the most prominent example of the expansive role a judge might play in settling class action litigation, its pitfalls and possibilities.

Judge Weinstein is perhaps most famous for his involvement in mass tort cases.[231] His approach has been to expand the litigation to embrace the entire problem before him through an expansive reading of procedural rules. This is in contrast to judges who traditionally use procedural rules to exclude parties, limit the issues and narrow or even dispose of the litigation before them.[232] Judge Weinstein's approach increases the participation of various involved groups in order to create consensus, while at the same time putting pressure on the parties through various legal rulings. One tool Judge Weinstein has used in pursuit of his goal of global settlement was community-wide consultation.[233] This practice

---

227. *See id.* (discussing drawbacks of expertise model in context of attorney leadership of civil rights movement).

228. As Minow explains: "Functionally, court-supervised settlements that establish systems for processing individual claims create temporary administrative agencies without proceeding through the legislative or executive branches." Minow, *supra* note 217, at 2020. Judicial action of this type may "trigger action by the other branches, and thereby promote the vision of overlapping and checking branches of government that lies behind the separation of powers." *Id.* at 2023.

229. *See* Rubenstein, *supra* note 14, at 1663 (discussing benefits of expertise model in context of attorney leadership of civil rights movement).

230. This is analogous to guardian *ad-litem* proposals. *See, e.g.*, Eric D. Green, *Advancing Individual Rights Through Group Justice*, 30 U.C. DAVIS L. REV. 791 (1997).

231. *See generally* WEINSTEIN, *supra* note 185; PETER H. SCHUCK, AGENT ORANGE ON TRIAL: MASS TOXIC DISASTERS IN THE COURTS (1987).

232. Minow, *supra* note 217, at 2013.

233. *See, e.g.*, Hart v. Cmty. Sch. Bd., 383 F. Supp. 699, 756-58 (E.D.N.Y.), *supplemented by* 383 F. Supp. 769 (E.D.N.Y. 1974), *aff'd by* 512 F.2d 37 (2d Cir. 1975) (involving not only school and parents but also teachers union, religious and other community leaders and experts in

is not as rare as it might seem.[234] Judge Weinstein also used his judicial power to consolidate litigation, sometimes through unreviewable preliminary orders.[235] He relied on a creative interpretation of the All Writs Act to transfer cases to his district, a legal strategy no longer available after *Syngenta v. Henson*.[236] These creative rulings, although not part of the traditional judicial role, do not depart from the traditional, dictatorial approach to class action governance.[237] They are an example of a "legal world created almost out of whole cloth by an individual judge."[238]

Despite some concerns, scholars and prominent class action activists have advocated increased judicial oversight as the solution to principal-agent problems, conflicts of interests and lack of fairness in class actions. These authors do not view the problem with class actions as a crisis of governance but as a failure of judicial oversight. They recommend a series of changes to Rule 23 to give judges additional criteria for approving settlements.[239] Similarly, other

---

various substantive areas in the resolution of civil rights class action).

234. In the *Mexico Money Transfer Litigation*, for example, the district judge heard evidence from leaders in the Mexican-American community as well as elected representatives. *See In re* Mexico Money Transfer Litig., 164 F. Supp. 2d 1002 (N.D. Ill. 2000), discussed *supra* note 20.

235. Through this mechanism, Judge Weinstein "eliminated" such intractable problems as the various state law claims of class members in the Agent Orange Litigation. Judge Weinstein "vaporized the choice-of-law problem" in the Agent Orange case by issuing a preliminary order that any state court would look to "national consensus law" to determine the manufacturers' liability, defenses and damages, and therefore the federal court would apply that same national consensus law to all the class members. This creative provisionary order was insulated from appellate review. *See* SCHUCK, *supra* note 231, at 128-31.

236. *See* Syngenta Crop Prot. Inc. v. Henson, 537 U.S. 28, 34 (2002) (holding that All Writs Act could not furnish removal jurisdiction and removal of cases was prohibited unless the movant could show original federal subject matter jurisdiction, even if petitioners were seeking to remove a case from state court to prevent the frustration of a federal court order).

237. "Outside the sphere of the adversary process and the rule-bound trial system, the settlement process permits room for personal persuasion, input, or pressure from the judge." Minow, *supra* note 217, at 2028. *See* SCHUCK, *supra* note 231, at 158-59 (describing the pressure Judge Weinstein put to bear on the parties in Agent Orange and the judge's role in dictating the settlement amount—one *lower* than what the parties had agreed to in that case). In discussing this pressure and its propriety, Minow notes "the exposure of the sheer will of the judge on the amount of the settlement works to remind all observers that law, and judging, inevitably reveal and express the views of specific, real persons." Minow, *supra* note 217, at 2029.

238. Minow, *supra* note 217, at 2030.

239. For example, Wolfman & Morrison, *supra* note 5, propose that Rule 23 should "require the court to reject settlements which provide no compensation for claimants who are giving up potentially viable claims, unless the court finds that the settlement provides benefits to those claimants that are comparable to those claims being abrogated." *Id.* at 498. They recommend that claims adjudicators should be allowed to take differences in state law and settlement history in various states into account, *see id.* at 499, advocate special scrutiny of non-monetary relief, and propose a new test for non-monetary settlements (e.g. coupons) requiring the deciding court to

scholars advocate guardians *ad-litem* as a prescription to governance problems, hoping that the installation of additional monitoring agents will strengthen judicial oversight.[240] These proposals, while useful, are incomplete solutions because they only address the role of the judge, ignoring the roles of other actors in the class action and the extent to which settlements are the results of the interactions of all these players.

III. A PRINCIPLED APPROACH TO GOVERNANCE

Class action governance is the relationship among various participants in determining the direction and outcome of the class action lawsuit.[241] Fundamental principles of good governance should control the relationships between the class members, the class counsel, objectors and the judges who oversee the suit and settlement. The concrete mechanisms of governance—the specific rules that control the unfolding of these relationships over time—should be designed to realize these broad principles.

No one analogy of governance fits the class action context perfectly. The corporate analogy is not entirely appropriate because claims are not fungible; in other words, there is no market—efficient or otherwise—for class action settlements that will mitigate the separation between ownership of claims and their adjudication. The analogy to democratic processes is incomplete because we are not willing to invest the resources required to obtain genuine direct participation in small claims class actions. The administrative analogy, while in many ways the closest to actual class action practice, does not address the misalignment of interests between class members and their counsel. As described above, the governance structure that has developed, dictatorship by class counsel with more or less judicial oversight depending on highly varying judicial practice, is not satisfactory nor is it required by Rule 23.

Some basic principles are deducible from the provisions of Rule 23 and due process jurisprudence, in conjunction with established principles of governance from the political, administrative and corporate contexts. The fundamental principles proposed here rely on the basic framework of existing due process jurisprudence and the newly revised Rule 23. Within that framework, these principles provide a coherent and systematic legal architecture that takes into account the special problems of class actions articulated in Part I.

Criticisms of previous proposals can also provide some guidelines for principles for class action governance. Market-based incentives take an overly

---

determine whether the non-monetary relief "provides all or substantially all of the class members a realistic opportunity to obtain valuable relief." *Id.* at 502. Finally, they make substantial suggestions concerning attorneys' fees. *See id.* at 503-07.

240. *See, e.g.*, Koniak & Cohen, *supra* note 9 (advocating greater use of guardian *ad litem* to monitor plaintiffs' attorneys' performance).

241. *See* ROBERT A.G. MONKS & NELL MINOW, CORPORATE GOVERNANCE 1 (2d ed. 2001) (defining corporate governance as "the relationship among various participants in determining the direction and performance of corporations").

simplistic view of the class action. For example, market mechanisms fail to address other structural factors that go into decision-making, such as the fact that class actions are a single unique transaction for claimants, and subsequent market penalties for singular instances of mismanagement are inadequate.[242] In addition, market mechanisms suffer from the persistence of incentives that encourage attorney collusion and rent seeking without passing on benefits to the class. Finally, creating a market in class actions would likely have prohibitive transactions costs. A robust system of governance needs to account for these structural factors. By contrast, mechanisms relying entirely on popular participation are too expensive to be worth the candle for small claims. Therefore, governance cannot be based on participation. Finally, mechanisms that focus on judicial oversight without attention to attorney incentives, objectors, and other participants, rely too much on judicial discretion by placing nearly the entire burden of process on judges, who may not be institutionally competent because of time constraints, lack of expertise, and lack of desire to construct a complete governance system on their own.

While courts have put much thought into evaluating substantive fairness of class actions, the procedures for ensuring fairness are inadequate.[243] To fill this gap, it is necessary to focus on the principles that should dictate the specific procedural mechanisms employed in class actions. By contrast, corporate governance regimes emphasize procedural governance guidelines, such as voting mechanisms, independent boards of directors and disclosure rules, while courts defer to management on substantive evaluations of corporate business decisions.[244] In the near term, a deferential standard, such as the business

---

242. *See* EASTERBROOK & FISCHEL, *supra* note 149, at 103 (discussing the ineffectiveness of market mechanisms on singular breaches of fiduciary duty in the corporate context).

243. For example, while courts will look at nine factors to determine the substantive fairness of a settlement, they will only look at four procedural factors, only one of which address the real problems in class action governance. *See* City of Detroit v. Grinnell Corp., 495 F.2d 448 (2d Cir. 1974) (articulating nine factors for determining whether a settlement is fair). The nine factors are:

> (1) the complexity, expense and likely duration of the litigation; (2) the reaction of the class to the settlement; (3) the stage of the proceedings and the amount of discovery completed; (4) the risks of establishing liability; (5) the risks of establishing damages; (6) the risks of maintaining the class action through the trial; (7) the ability of the defendants to withstand a greater judgment; (8) the range of reasonableness of the settlement fund in light of the best possible recovery; (9) the range of reasonableness of the settlement fund to a possible recovery in light of all the attendant risks of litigation. . . .

Id. at 463 (citations omitted); Girsh v. Jepson, 521 F.2d 153, 157 (3d Cir. 1975) (approving of nine *Grinnell* factors but remanding settlement because objectors were not given full and fair opportunity to be heard). The four procedural factors are discussed *supra* note 87 and accompanying text.

244. *See, e.g.*, Unocal Corp. v. Mesa Petroleum Co., 493 A.2d 946, 953-54 (Del. 1985).

> The business judgment rule is a "presumption that in making a business decision, the directors of a corporation acted on an informed basis, in good faith and in the honest belief that the action taken was in the best interests of the company." A hallmark of the

judgment rule, is not appropriate in the class action context. Even if an adequate governance regime were in place, judges cannot presume that class counsel is acting in the best interests of the class when the class action is a single transaction not susceptible to correction by market forces. While governance mechanisms alone cannot assure fairness, healthy governance when combined with a fairness inquiry is the most likely avenue to a robust and reliable class action system.

Procedural rules and standards that are well conceived are more likely to lead to consistent results, build realistic expectations, and reinforce a culture of compliance. There are several benefits to a principle-oriented procedural framework. It is comprehensive because rather than providing limited ad hoc solutions to isolated problems, procedural mechanisms and rules are designed to address all the relationships within the system. It also offers flexibility to both judges and class counsel and therefore encourages creativity in finding the right procedural mix. Finally, it is responsive to the changing landscape of the class action and permits the judge and attorneys to add procedural mechanisms where principles are not being met.

The four fundamental principles of class action settlement governance are (i) maximum disclosure, (ii) an actively adversarial process, (iii) expertise of decisionmakers, and (iv) independence of decisionmakers from influence and self-interest. The disclosure principle is integral both to our political structure and to corporation law. The requirement of an actively adversarial process is unique to the class action context. The principles of expertise and independence of decisionmakers are analogous to similar requirements of boards of directors and are familiar from both the political and corporate context. These four principles encompass most related principles. For example, accountability might be an obvious choice for a principle of class action governance. A combination of strict disclosure rules with an actively adversarial process will result in accountability of class counsel, so an accountability principle is arguably not analytically distinct from these fundamental principles.

The principled approach to governance requires judicial implementation of mechanisms to fulfill all four principles. A complete theory and practice of governance would fulfill these principles at each stage of the litigation. This proposal is confined to settlement, the most contentious moment in the class action suit. The mechanisms themselves may be altered, expanded, and/or contracted, while the principles remain static. Thus, these fundamental principles might be likened to governance guidelines and codes of best practices developed in the corporate sphere.[245] Firms can use any number of voting

---

business judgment rule is that a court will not substitute its judgment for that of the board if the latter's decision can be "attributed to any rational business purpose."

Id. at 954 (citations omitted).

245. *See* Holly J. Gregory, *Overview of Corporate Governance Guidelines and Codes of Best Practice in Developing and Emerging Markets*, *in* MONKS & MINOW, *supra* note 241, at 439. Unlike the governance guidelines discussed by Gregory, meeting the principles articulated herein would be mandatory though the specific mechanisms for realizing them would not have to be

mechanisms, appointment strategies and other procedural mechanisms to meet governance guidelines. The same is true of this class action governance regime.

A. Strong Disclosure Requirements

The first and perhaps most important principle for class action governance is the disclosure principle. This principle requires that material information be disclosed to class members, objectors, and judges. The principle of mandatory disclosure is necessary for good class action governance because disclosure of information alters the balance of power, decreases agency costs by enabling monitoring, and has a sanitizing effect.

Information is power. In the current structure of the class action, control over information resides almost exclusively with defendants and class counsel. Class counsel retains expert knowledge about litigation decisions, the substance and viability of claims, discovery, offers, terms of settlements, attorneys' fees, and side-deals. Control over this information provides class counsel with significant leverage over class members and potential objectors. By the manner in which class counsel releases and presents this information, it can control the relationship of class members and objectors to the class action. Thus, class counsel's information is a mode of control.[246] The distribution of information mirrors and reinforces the existing power structure, that is, the near total control over the litigation by class counsel with minimal oversight. This prevents challenges to the existing order because the lack of information available to class members, and especially objectors and judges, limits the ability of these important actors to challenge the existing governance structure.[247]

Mandatory disclosure would alter this imbalance, decentralizing power by preventing the asymmetry of information that currently characterizes class actions. The disclosure principle recognizes that the exercise of power within the

---

mandatory. There may be significant benefits to making specific, simple, rule-like mechanisms mandatory, as suggested by prominent corporations law scholars in a different context. *See* Bernard Black & Reinier Kraakman, *A Self-Enforcing Model of Corporate Law*, 109 HARV. L. REV. 1911, 1916 (1996).

246. *See generally* MICHEL FOUCAULT, DISCIPLINE AND PUNISH: THE BIRTH OF THE PRISON (1995) (presenting the now widely recognized theory that modern societies exercise power through the collection and use of information, and use this information and surveillance to internalize norms and values so that they become normalized). Foucault's theory of power relationships as discursive, that is, that power is created through discourse and the internalization of categories, has been utilized by scholars of privacy law. *See, e.g.*, Daniel J. Solove, *Privacy and Power: Computer Databases and Metaphores for Information Privacy*, 53 STAN. L. REV. 1393, 1418 (2001).

247. *See* Owen M. Fiss, *Why the State?*, 100 HARV. L. REV. 781, 786-87 (1987) (arguing that regulation of information, whether by the government or new agencies, limits political debate); *see also* Yochai Benkler, *Free as the Air to Common Use: First Amendment Constraints on Enclosure of the Public Domain*, 74 N.Y.U. L. REV. 354, 355 (1999) (arguing that the increase in privatization of information reinforces the existing power structure and hinders society's information production and exchange process).

class action is relational. Settlements are not produced solely by negotiation between class counsel and defendants' counsel, but by an interaction of a range of actors, including the attorneys, the court, opt-outs, and objectors.[248] Disbursing information to all of these actors will destabilize the controlling role of class counsel and empower other actors to contribute their important voices. Decisions about where information goes and how much will be released will determine the extent to which power is shifted to other actors. In this regard, the disclosure principle mandates maximum disclosure.

Second, disclosure decreases agency costs.[249] The most important way that mandatory disclosure reduces agency costs is by enabling informational intermediaries to monitor class counsel.[250] In the class action, informational intermediaries may include objectors, who analyze the information provided by class counsel and determine whether or not intervention and opposition to the appointment of counsel or to proposed settlements is appropriate; independent third parties appointed by the court to review settlement proposals or to follow up on the ultimate distribution of the settlement;[251] consumer advocates looking to prevent abuses;[252] and other attorneys looking to set up rival class actions.

Mandatory disclosure compensates for the absence of market competition to measure class actions, and for lack of an ongoing or multi-transactional relationship between class members and class counsel. Sophisticated scholars

---

248. This argument is similar to that articulated by scholars who write about intermediary organizations in the context of administrative and public law. These authors argue that administrative rules and regulations, as well as public services, are produced by interdependent networks of public and private partnerships rather than by top-down hierarchies. *See, e.g.*, Matthew Diller, *The Revolution in Welfare Administration: Rules, Discretion, and Entrepreneurial Government*, 75 N.Y.U. L. REV. 1121 (2000) (advocating new forms of accountability and arguing that by moving administrative agencies towards an entrepreneurial, incentive-based model, administrators have rendered decision-making less visible); Jody Freeman, *Extending Public Law Norms Through Privatization*, 116 HARV. L. REV. 1285 (2003) (critiquing the idea of "privatization" and arguing that public-private partnerships also extend public values to private actors); Susan Sturm, *Lawyers and the Practice of Workplace Equity*, 2002 WIS. L. REV. 277 (arguing in favor of a problem solving approach that harnesses and recognizes the role of intermediaries in creating norms in the workplace, rather than a top-down code of conduct approach).

249. *See* William M. Sage, *Regulating Through Information: Disclosure Laws and American Health Care*, 99 COLUM. L. REV. 1701, 1747-57 (1999) (explaining the uses of mandatory disclosure to reduce agency costs and deter agency failure *ex ante*).

250. On the use of informational intermediaries generally, see EASTERBROOK & FISCHEL, *supra* note 149, at 292-93. Examples of informational intermediaries in the corporate sphere include underwriters who price stock and auditors who evaluate firms' books.

251. This is similar to instances where class counsel does not take their fee until after the actual compensation has been paid out. In those cases counsel's fee is a percentage of the *actual* settlement amount, not of a projected amount.

252. Public Citizen Litigation Group and Trial Lawyers for Public Justice are examples of such groups.

in the corporate context have argued that disclosure is not empirically proven to prevent fraud or reduce agency costs.[253] Certainly, mandatory disclosure rules did not prevent the recent corporate graft perpetrated by the managers of companies such as Enron,[254] WorldCom,[255] Tyco,[256] and Qwest.[257] And because disclosure requires collecting and disseminating information, it is costly.

Even without empirical evidence as to the necessity of disclosure in the context of public corporations, the current disclosure regime for class actions seems merely to create more opportunities for self-dealing. The class members are essentially trapped in the class action. Unlike the management of a corporation, class counsel need not be concerned about attracting additional claimants through good disclosure practices or by proving their reliability. Class counsel need not devise a settlement that treats the class fairly, but merely a settlement not so egregious as to alert the judiciary or activists to its flaws. Class counsel need only be concerned with attrition so significant that it kills a settlement. Such a boycott is unlikely, and even if it were not, judges are generally inclined to approve settlements despite significant numbers of opt outs and objectors.[258] Thus, strict disclosure requirements compensate for the absence of the continuing, long-term relationship between class members and class counsel by requiring disclosure of information that an individual client would be able to obtain more easily.[259]

Third, mandatory disclosure works as an external discipline to increase

---

253. *See* EASTERBROOK & FISCHEL, *supra* note 149, at 287-97 (arguing that market mechanisms are just as or more likely to create a verification regime to attract capital). *But see* Merritt B. Fox, *Required Disclosure and Corporate Governance*, 62 LAW & CONTEM. PROBS. 113 (1999) (arguing that mandatory disclosure is necessary to educate shareholder-voters and assist shareholders to enforce management's fiduciary duties); Louis Lowenstein, *Financial Transparency and Corporate Governance: You Manage What You Measure*, 96 COLUM. L. REV. 1335 (1996) (arguing that mandatory disclosure makes management seek out information it might not otherwise obtain and therefore improve its management strategies).

254. *See* Paul Beckett et al., *Power Outage: How Energy Traders Turned Bonanza Into an Epic Bust—Unleashed by Deregulation, Industry Greed and Deceit Undid the Nascent Market*, WALL ST. J., Dec. 31, 2002, at A1.

255. *See* Rebecca Blumenstein & Susan Pulliam, *Leading the News: WorldCom Fraud Was Widespread—Ebbers, Many Executives Conspired to Falsify Results in Late 1990s, Probes Find*, WALL ST. J., June 10, 2003, at A3.

256. *See* David Armstrong, *Tyco to Restate Financial Results*, WALL ST. J., June 17, 2003, at A2.

257. *See* Dennis K. Berman & Deborah Solomon, *Ex-Executives Are Indicted in Qwest Probe*, WALL ST. J., Feb. 26, 2003, at B1.

258. Settlements with significant opt outs are often approved. *See supra* note 73 and accompanying text.

259. *See* Lowenstein, *supra* note 253, at 1344-45 (arguing analogously that disclosure compensates for the absence of long term, knowledgeable shareholders able to sit on corporate boards to represent other shareholders).

internal discipline.[260] Class counsel may behave differently when the prospect of transparency looms over them, and thus results may be improved *ex ante*. If sunlight is the best disinfectant,[261] then the sanitizing effect of mandatory disclosure may work from the inside as well as out—not just catching irresponsible behavior but encouraging better behavior to make disclosure less painful. The other side of this coin is that the absence of standardized disclosure requirements sends a message that hiding and manipulating information is acceptable or at least free of consequences.

Like sunshine laws in the political sphere, disclosure in the class action context builds trust in the settlement process by exposing it to public view. Such trust-building is especially beneficial to the class action bar which has received substantial criticism.[262] Unquestionably, disclosure requirements create additional monitoring costs for courts as well as costs for class counsel who must compile materials. There is a trade-off between the costs of complying with disclosure rules and the losses that classes and the public would suffer in the absence of disclosure. Whether disclosure rules will in fact produce social gains in excess of their cost is a matter for empirical study.[263] Even if disclosure does not create greater efficiencies from a market perspective, it is nonetheless a valuable counterpoint to the very serious criticisms of manipulation and lack of accountability that plague class action litigation.

To some extent, every rule regime will encourage potential violators to seek the edges of the rules or to manipulate the rules in order to obtain the best individual result. Moreover, the possibility exists that mandatory disclosure may have unintended negative consequences. For example, a vigorous disclosure regime may chill objectors who have little funding and for whom the expense of disclosure may be too high. Or disclosure may end settlement negotiations too early or give unfair advantage or information to defendants. There are also third-party problems with mandatory disclosure, particularly at the negotiation stage of the class action. Class counsel should not be required to release information that would hurt its ability to negotiate the best settlement for the class. Arguably, it is difficult to determine in advance what information should be privileged.[264]

For disclosure to be effective, it must be institutionalized properly. To be meaningful, disclosure must be comprehensive, comprehensible, and must happen at the right time in the decision-making process. Because of the potential for manipulation and the need to protect class counsel's ability to negotiate, the

---

260. For an interesting discussion of this idea in the corporate context, see *id.* at 1358.

261. LOUIS D. BRANDEIS, OTHER PEOPLE'S MONEY AND HOW BANKS USE IT 92 (1914) ("Sunlight is said to be the best of disinfectants; electric light the most efficient police man.").

262. *See supra* notes 9, 31 (describing criticisms of plaintiff's side class action attorneys).

263. If class counsel and defendants are required to release information regarding settlement outcomes after approval, it may be possible to do the empirical work to determine which procedural rules, if any, are making a difference in outcomes.

264. *See* Mars Steel Corp. v. Cont'l Ill. Nat'l Bank & Trust Co., 834 F.2d 677, 684 (7th Cir. 1987) (Posner, J.) ("Discovery of settlement negotiations in ongoing litigation is unusual because it would give a party information about an opponent's strategy.").

question of what information should be disclosed is as important as whether information is disclosed at all. A narrow disclosure standard might include only information material to the decision of class members and judges concerning the substantive fairness of the settlement.[265] A broad standard would require disclosure of all information related to the settlement and not subject to privilege, including side-settlements and agreements not currently subject to discovery, without any requirement of a judicial finding of materiality.[266] A third alternative would provide a two step process: a mandatory rule requiring the disclosure of certain specific information, in conjunction with a more liberal discovery rule that would permit discovery of information related to the settlement but not covered by privilege. Any of these alternatives would be superior to the current regime, but the most beneficial would be the most open standard, requiring disclosure of all information not subject to privilege.

Comprehensibility of disclosure depends on the audience for the information. Disclosures in notices sent out to the class directly would differ from disclosures to the court, objectors, or the public. In evaluating whether the disclosure principle is met, courts must consider audience, substance, and form, i.e., both the information disclosed and the manner in which it is disclosed. The disclosure principle cannot be met when attorneys provide information in ways that are difficult for the recipient to access or understand.

In notices to class members, there is a tension between completeness and comprehensibility. In complex settlements, notices should err on the side of comprehensibility.[267] It is the court's obligation to make notices understandable for the average citizen. The court should determine whether or not the notices are easily understood, both by reviewing them and by testing the notices on a representative group of class members prior to the actual distribution of the notices.

In disclosures to objectors' attorneys and sophisticated third parties, the court's oversight will relate to accessibility. Pleadings, motions, and orders are difficult for objectors to access because the courts where they are filed are often far away and copying court documents is expensive.[268] A solution to these access

---

265. This might echo the standard in the area of securities regulation, requiring disclosure of all information where there is a "substantial likelihood that the disclosure of the omitted fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available." Basic Inc. v. Levinson, 485 U.S. 224, 231-32 (1988) (quoting TSC Indus. Inc. v. Northway, Inc., 426 U.S. 438, 449 (1976)).

266. *See supra* text accompanying notes 85-86 (describing barriers to objector participation).

267. The notice could provide a two to three page summary of the settlement and make additional information available to interested claimants. Information can also be provided to claimants in a question and answer format which may be more accessible. Disclosures in the notices sent out to class members should be in plain English and in sufficiently large font for the average person to read. Many notices in class actions are so densely written as to be unintelligible even to a well-educated attorney. In addition, these notices are often in ten point font or less and are difficult to read.

268. In my experience, copying filed documents costs from $.25 to $.50 per page.

problems would be to require class counsel to post all notices, motions, opinions, orders, and other filings on a public Internet site.[269] This way, all information disclosed to the court and necessary for objections would be easily and cheaply accessible to objectors' counsel and interested claimants.

Disclosures should be required both before and after settlement approval. The minimum information necessary for disclosure prior to settlement approval includes (i) the total amount of attorneys' fees sought and method of calculation for those fees,[270] (ii) the attorneys' actual expenditure of hours and costs, (iii) any agreements between class counsel concerning labor and fees, (iv) predicted administrative costs for distribution of the fund, (v) the total amount of the settlement fund, (vi) the claimant response rate for similar class actions, (vii) the mean recovery per claimant or the predicted value of any coupon settlement on a secondary market, (viii) the content of any agreement concerning the distribution of excess settlement funds, and (ix) the content of any side agreements between class counsel or defendants and objectors or attorneys filing other class actions arising out of the same conduct by defendant, including the specific amount of any payments. This information should be disclosed prior to the filing deadline for any objections and prior to the Rule 23(e) hearing.

After settlement approval, class counsel and defendants should disclose (i) the total amount of attorneys' fees sought and received, (ii) the total amount of actual administration costs (such as the costs of cutting checks), (iii) the average amount of recovery per claimant who responded, (iii) the number of claimants responding to settlement, (iv) the total settlement fund actually paid to class members, (v) the number of opt outs, and (vi) if applicable, the actual value of any coupon settlement on a secondary market. The revised federal rules help somewhat in this regard by requiring courts to approve awards of attorneys' fees only after submission of a motion and requiring courts to make factual findings and conclusions of law regarding that motion.[271]

Post settlement reports, made under penalty of perjury, will provide policy makers, judges, and objectors a means for evaluating and comparing settlements. It also gives both sides an incentive to make sure that their rosy predictions at the

---

269. Telephone Interview with Alan Morrison, Director, Public Citizen Litigation Group, Washington, D.C. (July 15, 2003).

270. *See* Staton v. Boeing Co., 313 F.3d 447 (9th Cir. 2002) (holding that notice of amount of attorneys' fees is required), *opinion withdrawn and superseded on denial of reh'g by* 327 F.3d 938, 963 n.15 (9th Cir. 2003) (determining to "scrutinize the attorneys' fees provisions with special care . . . [because] class notice did not break out the amount of attorneys' fees provided for in the settlement agreement . . . ."); Goldenberg v. Marriott PLP Corp., 33 F. Supp. 2d 434, 441 (D. Md. 1998) ("Notice of the potential extent of attorneys fee awards is deemed essential because it allows class members to determine the possible influence of the fees on the settlement and to make informed decisions about their right to challenge the fee award.").

271. *See* FED. R. CIV. P. 23(h). This brings class action fee issues closer to the rules governing civil rights fee-shifting cases, where courts decide what amount of attorneys' fees are reasonable pursuant to motion. *See* 42 U.S.C. § 1988 (2000).

fairness hearing are realistic.[272] The limited empirical evidence available on this question indicates that self-reporting correlates with superior outcomes.[273] Some class action firms voluntarily provide clauses in stipulations of settlement requiring that some or all of the attorneys' fees be paid after the class has been paid and that the fees reflect the total payout.[274] But such disclosure cannot depend on the occasional voluntary act of progressive class counsel.

Objectors play a pivotal role as informational intermediaries. A robust disclosure principle would provide objectors and potential objectors an opportunity to obtain information about all aspects of settlement, including side deals between the participants in the litigation and third parties. Currently, to obtain discovery, objectors must show a reason to believe that the side deal affected the settlement.[275] Objectors are unlikely to meet this burden, and given the monopoly class counsel has over the class action and the misalignment of interests, side settlements are inherently suspect. Likewise, objectors who make side deals to drop objections must be required to publicly disclose the terms of all such settlements and to have them approved by the court.

The involvement of these intermediaries raises "superagency" problems, that

---

272. *See Staton*, 327 F.3d at 958 n .12.

> Even when there is no direct proof of explicit collusion, there is always the possibility in class action settlements that the defendant, class counsel, and class representatives will all pursue their own interests at the expense of the class. For that reason, the absence of direct proof of collusion does not reduce the need for careful review of the fairness of the settlement, particularly those aspects of the settlement that could constitute inducements to the participants in the negotiation to forego pursuit of class interests.

Id.

273. For example, in *Pinny v. Great Western Bank*, No. CV 95-2110 (C.D. Cal. 1997), a case study described in depth in the 1999 RAND study, class counsel provided a final report showing the actual settlement paid out, which was nearly identical to what had been predicted. *See* HENSLER ET AL., *supra* note 2, at 175-84. By comparison, in *Roberts v. Bausch & Lomb, Inc.*, also discussed in HENSLER ET AL., *supra* note 2, at 145-73, the total settlement was presented to the court as approximately $68 million depending on how many class members claimed rewards. The attorneys were paid $8 million based on that representation, calculating a fee of a little less than fifteen percent. The parties were not required to report the ultimate payout to the court, but based on SEC filings RAND researchers deduced that the defendant never allocated more than $37.7 million to pay out all expenses relating to the litigation, including attorneys' fees and administration costs. Thus, there was likely a substantial overpayment to class counsel.

274. *See, e.g.*, Submission of Proposed Stipulation of Settlement, Tyson v. City of New York, 97 Civ. 3762 (JSM) (S.D.N.Y., Jan. 5, 2000) (requiring that fees to class counsel be judicially approved and not paid out until administration of class action was completed) (on file with author).

275. *See* Duhaime v. John Hancock Mut. Life Ins. Co., 183 F.3d 1 (1st Cir. 1999). In *Duhaime*, the First Circuit held that objectors could not have access to the substance of a side-settlement between objecting class members who had appealed settlement approval. *Id.* at 4. Non-appealing objectors were denied discovery concerning this side settlement because the court found that the side settlement did not "affect" the main settlement. *Id.*

is, who will monitor the intermediaries?[276] The natural answer is judges. Class counsel and defendants complain that many objectors are fee-seeking mercenaries attempting to hold settlements hostage. Objectors' counsel should be held to the same standard as class counsel, but they should not be held to a higher standard that will chill objections.[277] Objector intervention will likely increase when objectors' counsel are encouraged by monetary incentives such as the award of attorneys' fees or buy-outs to preserve global peace. Nevertheless, the fear that some attorneys will take advantage of the system, whether they are serving as class counsel or as objectors' counsel, should not prevent courts from encouraging objectors to air the limits or negative aspects of settlement.

B. An Actively Adversarial Process

The adversarial principle requires that the formulation and evaluation of settlements be reached through a vigorous, active adversarial process. This principle recognizes two factors in class action practice. First, active class member participation is impossible and more than likely undesirable in Rule 23(b)(3) class actions. Second, in the absence of client monitoring and involvement, a vigorous adversarial process is necessary to expose inadequate settlements and produce optimal outcomes. Accordingly, rather than focusing on direct participation or class member activism, this principle takes a broader view and redirects our lens to defendants, class counsel, the court, objectors, and the extent to which their relationships can be structured to ensure that the concerns of absent class members are aired and addressed.

The requirement that settlements be reached through arms length negotiation,[278] is a poor solution to concerns that settlements are not currently reached through a sufficiently adversarial process that would expose all the competing interests and attempt to resolve them in the most equitable manner possible.[279] Even with arms length negotiation, class counsel's interests may

---

276. *See* Samuel Issacharoff & Daniel R. Ortiz, *Governing Through Intermediaries*, 85 VA. L. REV. 1627 (1999).

277. *See supra* note 100 and accompanying text (discussing chilling effect of sanctions on objectors).

278. *See supra* text accompanying note 87 (discussing requirements for procedural fairness in class action settlements: arms length negotiations, sufficient discovery, proponents were experienced, and limited objections).

279. As Judge Easterbrook so eloquently put it,

> Representative plaintiffs and their lawyers may be imperfect agents of the other class members—may even put one over on the court, in a staged performance. The lawyers support the settlement to get fees; the defendants support it to evade liability; the court can't vindicate the class's rights because the friendly presentation means that it lacks essential information.

Kamilewicz v. Bank of Boston Corp., 100 F.3d 1348, 1352 (7th Cir. 1996) (Easterbrook, J., dissenting from denial of rehearing en banc), *quoted with approval in* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 621(1997).

impede her ability to serve the class. Furthermore, once they reach settlement, defendants and class counsel share an interest in getting the settlement approved. Commentators have also lamented that many of the problems in settlement are the result of a lack of adversarial process, which in turn is the result of lack of client monitoring by class counsel.[280]

Democracy theorists have pointed to participation and deliberation as of one of the salient characteristics of a democratic society.[281] But direct and active class member participation is impossible because in consumer class actions participation is too expensive in relation to the interests at stake. For instance, democratic process might drain potential recovery to the extent that insufficient resources would be left to attract class counsel.[282] Furthermore, because the stakes for individual class members are low, the external investment in participation must be all the more significant to yield results. There are good utilitarian reasons for not investing substantial resources in pursuit of a direct democracy in class action governance.[283] As a society we do not want to invest the amount of resources necessary to render the class action a fully deliberative and participatory process in the way that we are willing to invest in our political process.

This does not mean, however, that there is no room for deliberation in the class action context. Nor does it mean that the only means of participation and communication in class actions need be passive, such as opt-out mechanisms.[284] Because of their deterrent and compensatory functions, small claims class actions

---

280. *See* Issacharoff, *supra* note 60, at 348 (arguing that even though the Supreme Court warned against the evils of settling without reference to testing of claims in the adversarial process in *Amchem*, that is just what settlement classes are).

281. *See* Michelman, *supra* note 158, at 1503 (discussing the importance of political participation as a positive human good).

282. *See generally* Coffee, *Understanding the Plaintiff's Attorney*, *supra* note 24, at 725 (discussing arguments in favor of substantial attorneys fees). For example, in addition to their time and other expenditures, class attorneys are required to pay for notice and conformity with other due process requirements. *See* Eisen v. Carlisle & Jacqueline, 417 U.S. 156, 177 (1974) (requiring that plaintiffs pay cost of notice). This rule creates an incentive towards settlement-only certifications, because once a settlement has been reached, defendants will often pay the costs of notice and other administrative costs.

283. In other contexts, there may be normative arguments in favor of expensive participation regimes. One could imagine a regime where civil rights class actions would require the creation of a truly deliberative process because of the special connection between civil rights laws and our larger democratic processes. In that case, too, fee shifting statutes require that the violator pay attorneys fees, and thus the additional expenditure on democratic process would not come at the expense of attorney incentives to bring suits. Because those are not the types of class actions discussed here, I leave the question of whether civil rights is a special case to another day.

284. *See supra* Part I.C (arguing that the opt out solution is ineffective because the small claims class action is essentially a non-competitive space during the critical process of settlement approval and noting that while opt outs may be more valuable during certification, class members are hampered by lack of information and the incentive towards reverse auctions).

are a public good. Engaging in a deliberative process in the courtroom is still the best method to probe issues the judge may consider in approving or rejecting settlements. As we saw earlier, decisionmakers will often only address those problems they have studied and measured.[285] An adversarial process will improve deliberation by expanding the depth and breadth of the issues brought before the court. Such a process may have the added benefit of helping to form class member preferences in favor of reasonable settlements.[286]

The size of settlement funds and manner in which these funds are distributed should be the subject of deliberation and discussion. Under the current regime, the class representative is an 'authorized' representative, meaning one to whom the court gives the power to authorize actions on behalf of the class and little more.[287] Courts do not look too deeply into the self-interest of class representatives to determine, for example, the effect of additional payments on the class representatives.[288] Nor do courts inquire whether class representatives were consulted in the formulation of the settlement.

Fairness hearings provide a perfect avenue for deliberation, but have historically tended to be insubstantial and pro-forma. The doctrines encouraging settlement and limiting review reinforce this reality. Some courts have expressly limited settlement hearings to prevent what they feared would become a full-blown trial on the merits, and fear that increased support of objectors will do just that.[289] Such courts take too narrow a view of the fairness hearing.[290] Rule 23(e) hearings are intended to protect class members. Class members have only limited means of communicating through (or with) counsel and class counsel have their own set of interests that are not aligned with those of class members. During settlement, class counsel and defendants share an interest in getting the settlement approved, regardless of its objective fairness. In order to expose unfairness, Rule 23(e) hearings should be vigorous and adversarial, involving a variety of agents, including class counsel, objectors' counsel, class members and third parties.[291]

---

285. *See* Lowenstein, *supra* note 253, at 1335.

286. *See* BARTELS, *supra* note 89, at 304 (positing that "actual outcome of political processes must be determined not by preferences alone but also by the structure of political institutions that channel preferences in particular ways").

287. *See supra* Part II.B.1 (discussing the concept of representation in the class context).

288. While such payments may be necessary to encourage class members to participate as representatives in small claims class actions, it also means that their interests are not aligned with those of the class.

289. "The temptation to convert a settlement hearing into a full trial on the merits must be resisted." Mars Steel Corp. v. Cont'l Ill. Nat'l Bank & Trust Co., 834 F.2d 677, 684 (7th Cir. 1987) (Posner, J.).

290. *See* Wolfman & Morrison, *supra* note 5, at 488 (describing a half-hour long fairness hearing).

291. For example, in *Hoffman v. BancBoston Mortgage Corp.*, No. 91-1880 (Ala. Cir. Ct. Jan. 24, 1994), a particularly egregious settlement which resulted in plaintiffs paying more in attorneys' fees than they received in the settlement, there was only one objector. That case is described in

The fairness hearing should be a substantive presentation of several viewpoints on the settlement. In this sense, the fairness hearing will be a kind of trial on the merits of the settlement, although not a trial on the underlying merits of the case.[292] Reconceiving the settlement hearing as a bench trial, or even a jury trial, is the working metaphor for the kind of adjudicatory procedure we should be looking for.[293]

For a Rule 23(e) hearing to be adversarial, it first requires adversaries, who may be unprompted objectors represented by counsel or objectors solicited by the court.[294] In the absence of self-motivated objectors represented by competent counsel, the adversarial principle requires that the court appoint a third party to act as a "devil's advocate" for the class, such as a guardian *ad litem*.[295] Such an appointed "objector" should bring her independent evaluation of the substance of the settlement, as well as its procedure, including the extent to which that procedure has met the four principles discussed here. She should also evaluate opt outs and determine what motivated any decision to opt out. A third-person evaluation of a settlement from claimant's point of view should expose problems glossed over in the presentation of the settlement by self-interested counsel.[296]

Another avenue for introducing intermediaries in the absence of vigorous objectors is to appoint advocacy groups such as Public Citizen Litigation Group

---

Koniak & Cohen, *supra* note 9, at 1058-74.

292. The current view, as discussed *supra* at 86, is that the Rule 23(e) hearing not resemble a trial at all.

293. *See* Koniak & Cohen, *supra* note 9, at 1129-30 (proposing settlements be heard by juries).

294. *See, e.g.*, Bell Atl. Corp. v. Bolger, 2 F.3d 1304 (3d Cir. 1993).

> In assessing settlements of representative actions, judges no longer have the full benefit of the adversarial process. . . . In seeking court approval of their settlement proposal, plaintiffs' attorneys' and defendants' interests coalesce and mutual interest may result in mutual indulgence. The parties can be expected to spotlight the proposal's strengths and slight its defects. In such circumstances, objectors play an important role by giving courts access to information on the settlement's merits.

Id. at 1310 (citations omitted).

295. *See, e.g.*, Alon Klement, *Who Should Guard the Guardians: A New Approach to Monitoring Class Action Lawyers*, 21 REV. LITIG. 25 (2002) (proposing private monitors who would be paid out of settlement fund). *But see* Koniak & Cohen, *supra* note 9, at 1111 (discussing problems faced by *guardians ad litem* and stating that "we can report without attribution, for whatever it may be worth, that the guardians we have talked to understand their job is to approve the deal that the settling parties have constructed, after suggesting a few minor changes, not to recommend that the settlement be chucked"). The concern about guardians and other objectors having their own self-interest is discussed below.

296. The value of such independent analysis can be best illustrated in Koniak & Cohen, *supra* note 9, in which these scholars provide a thorough and shocking analysis of a misleading settlement in which class members who received nothing were nevertheless required to pay substantial attorneys' fees which were deducted from their escrow accounts without their direct consent. One hopes that such a thorough presentation would have affected the court's decision in that case.

or Trial Lawyers for Public Justice as class guardians.[297] Not-for-profit groups are more likely to provide an independent evaluation of settlements and therefore to contribute to the adversarial nature of the process. Objectors' counsel seeking fees, or guardians chosen by the defendants and class counsel, by contrast, are more likely to be subject to influences that could prevent them from opposing an inadequate settlement. If they are independent, state attorneys general or consumer advocacy groups might also serve this intermediary role.[298]

In most settlements, objectors' counsel (including not-for-profits) need to be encouraged by the payment of fees, just as class counsel does. How much of this type of encouragement should courts provide? It is clear that some incentive is necessary as objectors are a critical part of the settlement process. Courts have held that any objectors who contribute materially to the proceeding may obtain a fee.[299] Some courts have interpreted a "material contribution" to include a contribution to the adversarial nature of the proceedings and the provision of representation for a significant group of objectors.[300] Other courts have adopted a much narrower approach, limiting the recovery of objectors' counsel's fee to where they "produce an improvement in the settlement worth more than the fee they are seeking; otherwise they have rendered no benefit to the class."[301] On an

---

297. These groups are described in HENSLER ET AL., *supra* note 2, at 89-90. The positive involvement of Trial Lawyers for Public Justice in one settlement is described in *id.* at 201-04.

298. *See* Koniak & Cohen, *supra* note 9, at 1083 (discussing proposal to mandate alerting state attorneys general of all class action settlements).

299. *See, e.g.*, Petrovic v. Amoco Oil Co., 200 F.3d 1140, 1156 (8th Cir. 1999) ("To recover fees from a common fund, attorneys must demonstrate that their services were of some benefit to the fund or enhanced the adversarial process."); Gottlieb v. Barry, 43 F.3d 474, 490-91 (10th Cir. 1994) (awarding fees to attorneys for objectors whose work resulted in "a reduction of certain fee and expense awards, and thereby benefited the class"); Fisher v. Procter & Gamble Mfg. Co., 613 F.2d 527, 547 (5th Cir. 1980) ("In opposing the consent decree in a case in which the plaintiffs ultimately prevailed at trial, the objectors benefited their class and helped to vindicate the important public rights protected by Title VII."); Lindy Bros. Builders, Inc. v. Am. Radiator, 540 F.2d 102, 112 (3d Cir. 1976) (attorneys fees awarded from a common fund depends on whether the attorneys' "specific services benefited the fund whether they tended to create, increase, protect or preserve the fund"); White v. Auerbach, 500 F.2d 822 (2d Cir. 1974).

> Accordingly, it is well settled that objectors have a valuable and important role to perform in preventing collusive or otherwise unfavorable settlements, and that, as the district court recognized, they are entitled to an allowance as compensation for attorneys' fees and expenses where a proper showing has been made that the settlement was improved as a result of their efforts.

Id. at 828.

300. *See* Elliott v. Sperry Rand Corp., 680 F.2d 1225, 1227 (8th Cir. 1982) (per curiam) (granting award of attorneys' fees to representatives of objectors who were formerly representatives, and ruling that amount of fees must come from fee set-aside in settlement not from general fund).

301. Reynolds v. Beneficial Nat'l Bank, 288 F.3d 277, 288 (7th Cir. 2002) (Posner, J.); *see also* Class Plaintiffs v. Jaffe & Schlesinger, P.A., 19 F.3d 1306, 1308-09 (9th Cir. 1994) (per curiam) (affirming denial of attorneys' fees for work of attorneys in state court class action settled

abstract level, the latter approach is more narrowly aimed at defeating irresponsible objectors. Just as fees for class counsel should be linked to the benefits they produce for the class, fees for objectors' counsel should be linked to the benefits they produce.[302] But a measure of objectors' counsel's fees that is linked too closely with a specific monetary increase in the settlement fund may encourage objectors to withdraw valid objections in order to settle with class counsel, especially when those objections are sufficiently serious to scuttle rather than increase a settlement. Accordingly, the adversarial principle requires a broad view of compensation for objectors' counsel. Under this view, attorneys who make a material contribution to the adversarial nature of the proceedings and assist in the realization of the other principles of governance should be compensated even if they did not produce a monetary improvement in the settlement.

Although opt outs, as currently exercised do not contribute to the adversarial process, some procedural adjustments may improve this situation. First, the timing of opt outs should be altered to permit opt outs *after* the 23(e) hearing rather than prior to the hearing as currently practiced. This would encourage objections by permitting potential opt outs to state their objections and influence the process prior to the adoption of a settlement. It would also allow individual claimants to take advantage of the information adduced prior to and during the fairness hearing in choosing whether or not to participate in the settlement.[303] Second, opt outs might be made a collective right rather than an individual right.[304] This would permit state attorneys general or other authorized persons to represent sub-groups for whom the settlement is unfair and may prevent some collateral attacks against settlements.[305] Since opt outs do not permit claimants to realize their autonomy, the only way they will constitute an expression of choice is if they are collectivized. Permitting collective opt outs is not so different from authorizing separate representation of sub-groups or the initial appointment of class counsel.[306] In both cases, the legitimacy of representation

---

in federal court class action because the work was attenuated and stating that "[w]e know of no authority which mandates an award of fees to attorneys not formally representing the class, whose activities in representing others incidentally benefit the class").

302. *See generally* Coffee, *Understanding the Plaintiffs' Attorney*, *supra* note 24 (arguing that fee awards in class actions should be a percentage of recovery rather than the lodestar approach).

303. *See, e.g.*, *In re* Brand Name Prescription Drugs Antitrust Litig., 115 F.3d 456 (7th Cir. 1997) (barring opt outs from appealing approval of settlement). It seems likely that the opt outs in that case were attempting to use the leverage of individual suits to extract rent from class counsel or defendants.

304. *See supra* note 68 (discussing opt out as an individual right).

305. *See* State v. Homeside Lending, Inc., 826 A.2d 997 (Vt. 2003) (holding that class action settlement did not preclude suit by Vermont residents because Alabama Court lacked personal jurisdiction due to defects in due process protections).

306. *See* Ortiz v. Fibreboard Corp., 527 U.S. 815, 856 (1999) (requiring sub-classification where diverse groups within class have conflicts of interests); *see also* FED. R. CIV. P. Rule 23(g) (requiring court appointment of class counsel).

is in question and the class action risks fragmentation.

A requirement that class counsel communicate with and build consensus among class members, as has been suggested by some commentators,[307] is insufficient to comply with the adversarial principle because class counsel is likely to learn from class members only what class counsel wants to know. Representatives might be included in the process by permitting objectors to cross-examine appointed representatives about their understanding of the settlement. In the alternative, an independent "jury" of representative claimants could be presented with the settlement to determine whether or not it is fair.[308] Such a jury could be made up of a random selection or a representative cross-section of class members and receive a payment commensurate with the work they do (rather than for approving the terms of a settlement and perhaps the award of attorneys fees). The same problems that plague jury trials would be a problem here, particularly the fear that the claimant jury would lack the requisite expertise or ability to master the issues at stake as well as a judge with substantial expertise in class action litigation and settlement. A jury may be as well placed as a judge to determine whether a settlement is substantively fair once they understand it. Furthermore, it may be more in keeping with our concept of settlements as a type of private ordering to allow a jury of claimants to make this determination. Even if we believe the court is better situated to make fairness determinations, a jury of claimants may be given a consultative role, including the ability to ask questions and express concerns to the judge.[309]

C. Expertise of Decisionmakers and Agents

Quality of leadership is a well-recognized principle of corporate governance that applies equally in the class action context.[310] Quality may mean different things for different decisionmakers. The three participants whose expertise is most important are class counsel, judges and objectors' counsel. With respect

---

307. *See* Rubenstein, *supra* note 14, at 1670-71 (proposing requirement of community consultation); Shapiro, *supra* note 21, at 959 (proposing representative group of claimants to curb excesses of class counsel).

308. These representatives should not be persons chosen by class counsel in order to preserve their impartiality.

309. *Cf.* Mariano-Florentino Cuellar, Rethinking Public Engagement in the Administrative State 70-74 (2003) (unpublished manuscript on file with author) (proposing negotiated rulemaking in which persons affected by a proposed rule are given an opportunity to be educated about it and have input with legal effect).

310. For example, Business Week declared "Director Quality" one of five good governance principles that investors should look for. *See* Louis Lavelle, *Special Report: The Best And Worst Boards: How The Corporate Scandals Are Sparking A Revolution In Governance*, BUSINESS WEEK, October 7, 2002 at 104 (defining "Director Quality" as follows: "Boards should include at least one independent director with experience in the company's core business and one who is the CEO of an equivalent-size company. Fully employed directors should sit on no more than four boards, retirees no more than seven. Each director should attend at least 75% of all meetings.").

to class counsel, quality has traditionally meant experience in litigating similar claims.[311] This principle is the one that has been most consistently addressed in the present class action regime, and is expressed in the newly adopted Rule 23(g), which requires the court to evaluate the experience and quality of counsel prior to appointment.

Quality of decision-making as applied to judges is a more difficult concept to discuss, because to raise this issue seems to impugn the intelligence, expertise, work ethic and level of commitment of judges. Already embedded into the structure of class action practice are mechanisms that recognize the importance of experience in complex litigation. For example, Multi-district Litigation panels are an attempt to consolidate decision-making with judges who have special expertise in class actions.[312] Appellate deference to district court judges' fairness determinations likewise recognizes the importance of expertise, in that case the district court's case-specific expertise.[313] Nevertheless, there are some indicia that case-specific expertise is not so important. For example, Multi-district Litigation panel judges do not retain cases referred to them for trial,[314] and appellate courts do not always defer to the case-specific expertise of district courts.[315]

Finally, objectors' counsel's history and experience, as well as their financial dealings with class counsel and defendants, should also be subject to judicial scrutiny. Standards for evaluating objector's counsel need not be different than those governing class counsel. Objectors represented by experienced counsel will likely be given more attention than objectors without representation.[316] This may mean that objectors lacking representation will need to have a guardian appointed on their behalf. Experienced not-for-profit objectors, such as Public Citizen Litigation Group, Trial Lawyers for Public Justice, or states' attorneys general should be encouraged to fill this role because of their limited financial interest and their history of responsible involvement in egregious cases.

---

311. The trend in Rule 23 jurisprudence is towards increasingly specialized substantive areas of class action practice, so that a consumer class action is sufficiently different from a mass tort class action to require separate expertise.

312. *See* 28 U.S.C. § 1407(a) (2003).

313. *See supra* notes 106, 108, 218 (discussing standard of review).

314. *See* Lexecon, Inc. v. Milberg, Weiss, Bershad, Hynes & Lerach, 523 U.S. 26 (1998) (holding that district court judge hearing case pursuant to multi-district litigation statute has no authority to transfer case to himself for trial). There has been significant criticism of this decision and Congress has considered bills to reverse this result. *See, e.g.*, HR 1756, 106th Cong. 1st Sess. (1999).

315. Arguably, in *Amchem Products, Inc. v. Windsor* the Supreme Court made an independent determination of the fairness of that settlement. 521 U.S. 591 (1997). *See* discussion accompanying *supra* note 218.

316. *See In re* Westinghouse Sec. Litig., 219 F. Supp.2d 657 (W.D. Penn. 2002) (denying attorney's fees award to pro se objector who benefited class).

D. Independence of Decisionmakers from Influence and Self-Interest

The final principle for class action governance is the independence principle. Independent decisionmakers are decisionmakers who are not likely to gain individually from the decisions they make, so that they are free from incentives towards self-dealing transactions. That class counsel are not independent because of their financial interest in the outcome of class action settlements is the staple of much of the literature on class actions.[317] As described in Part II, much of this literature focuses on the agent-principal problem and on the incentives influencing the behavior of the attorney as decisionmaker.[318] Fewer have focused on the role of the judge as decisionmaker.[319] Both kinds of decisionmakers are necessary to any class action governance regime and class members are forced to rely on their expertise. So too, state attorneys general, objectors' counsel and *guardians ad litem* are not always independent of self-interest. The integrity of all these decisionmakers is particularly of concern where, as in the class action context, they are not selected through an arms-length bargaining process or by popular vote, but by some more suspect mechanism: either judicial appointment or self-appointment.[320]

---

317. Some scholars characterize the attorney self-interest in settlement as an incentive based problem which is not the same as collusion or self-dealing. *See* Coffee, *Understanding the Plaintiffs' Attorney*, *supra* note 24. Other scholars see this problem as resulting in rampant breaches of fiduciary duty. *See* Koniak & Cohen, *supra* note 9. This difference is a matter of degree rather than of kind.

318. *See, e.g.*, Coffee, *Accountability*, *supra* note 3 (proposing a competing proxy-type solution to give attorneys incentive to provide better representation); Coffee, *Understanding the Plaintiffs' Attorney*, *supra* note 24 (advocating that attorneys' fees in class actions be a percentage of the fund rather than loadstart to create an incentive to maximize compensation); Macey & Miller, *supra* note 7 (advocating a regime of auctions for class counsel to encourage better quality representation); Fisch, *supra* note 120 (arguing in favor of increased reliance on stronger lead plaintiffs to control attorney excesses); Rhode, *supra* note 111 (arguing for greater ethical obligations for attorneys representing class to seek out class opinions); Rubenstein, *supra* note 14 (arguing, in part, in favor of an expertise model allowing attorneys to make litigation decisions); Susan P. Koniak & George M. Cohen, *In Hell There Will Be Lawyers Without Clients or Law*, 30 HOFSTRA L. REV. 129 (2001) (discussing the difficulties of the lawyer-client relationship in the class action context and the possibilities for abuse); Koniak & Cohen, *supra* note 9, at 1103-19 (arguing for the encouragement of subsequent suits to deter attorney misconduct and collusion in crafting settlements).

319. Two notable examples of articles that take this issue head on are Koniak & Cohen, *supra* note 9, at 1105, 1123-28 (describing the incentives judges have to approve collusive or inadequate settlements); Minow, *supra* note 217 (discussing the tensions created by judicial involvement in settlement making).

320. Proposed Rule 23(g) would require judicial appointment of class counsel. It is not clear, however, whether this change in the rule would make much of a difference from current practice. Koniak and Cohen make a very convincing argument that class action attorneys' use of side deals to remove competing attorneys from the playing field and retain control over the class action

With regards to class counsel, meeting the principle of independent decision-making requires two types of mechanisms, internal and external. The first type concerns internal incentive mechanisms, such as the compensation scheme for class counsel. Internal mechanisms might include an initial auction for services combined with compensation that should rise proportionately to actual claimant recovery, and judicial oversight of the fairness of that compensation in light of the ultimate outcome of the case.[321] The problem with pure percentage payments is that they have an arbitrary quality to them. What percentage will sufficiently encourage valid small claims class actions without giving attorneys a windfall out of class members' pockets? Judicial review of attorney compensation or the attorneys' fee motion should be accompanied by an evaluation either by objectors or other intermediaries.[322] This would mimic the corporate use of independent auditors to review filings. Generally, incentive mechanisms based on attorney compensation should be linked to actual benefits conferred rather than predicted outcomes and should bear some relationship to the effort expended by class counsel.[323] The linking of fees to the actual amount paid out in settlement, rather than counsel's predictions, should be mandatory.[324]

Undeniably, fraud or misleading representations can be made even with the protection of 'independent' audits and reporting requirements.[325] This type of malfeasance requires external control mechanisms. Corporate governance scholars have suggested that strong legal remedies with simple bright line rules will produce the best compliance results in an economy with limited resources for enforcement.[326] Such deterrents might include the encouragement of

---

constitutes a potential antitrust violation. *See* Koniak & Cohen, *supra* note 9, at 1200-06. As discussed in Part II, such auctions likely will not solve other problems associated with class representation.

321. Rule 23(h) now requires judicial approval of "reasonable" attorneys' fees pursuant to motion, much like civil rights practice.

322. This is to avoid situations in which class counsel bases its fee figure on a percentage of the total compensation fund but only a fraction of that fund is eventually paid out, leaving counsel with a windfall. *See* discussion *supra* accompanying notes 273-74. For an excellent description of the sophistication of these kinds of abuses, see Koniak & Cohen, *supra* note 9, at 1083-84.

323. The attorneys' fees in most of the egregious cases of poor settlements seem to be calculated based on predictions of benefit for the group rather than actual benefits conferred. *See* supra note 273 (discussing the Bausch & Lomb litigation); *see also* Koniak & Cohen, *supra* note 9, at 1058-74 (describing Hoffman v. BancBoston Mortgage Corp., No. 91-1880 (Ala. Cir. Ct. Jan. 24, 1994), in which the attorneys' fee calculation was based on moneys claimants already had and not on any benefits conferred by class counsel).

324. *See, e.g.*, *Tyson* class action, discussed *supra* note 274.

325. The role of the Arthur Anderson accounting firm in the most recent corporate scandals is an excellent example of this. *See* Kurt Eichenwald, *Enron's Many Strands: the Investigation. Anderson Charged with Obstruction in Enron Inquiry*, N.Y. TIMES, Mar. 15, 2002, at A1.

326. *See* Bernard Black & Reinier Kraakman, *A Self-Enforcing Model of Corporate Law*, 109 HARV. L. REV. 1911, 1916 (1996) (arguing in favor of bright line rules that are easy to follow accompanied by strong legal remedies on paper to compensate for low probability that these

malpractice suits to deter class action attorneys who take advantage of their client by abusing attorneys' fee allocations, or the use of disciplinary proceedings against them.[327] Scholars critical of the plaintiffs-side class action bar have pointed out that the cost of litigation under the current regime of attorney control make pursuing such suits very difficult for wronged plaintiffs,[328] and have noted that disciplinary regimes can be ineffectual.[329] This is especially so where sophisticated wrongdoing is unlikely to be captured by bright-line rules. In the context of allegations of attorney misconduct, there is a trade-off between bright-line rules and standards. While rules are easily administered, they are likely to be underinclusive, targeting egregious misconduct but missing more sophisticated misconduct. Standards can encompass more sophisticated misconduct, but are difficult to prosecute and may overdeter attorneys from bringing suits.[330] A bright line rule with strong legal remedies for breach is probably the most efficient choice. Even if such external mechanisms only deter the most blatant and egregious abuses, they nevertheless should not be ignored.

Independent and objective decision-making is also an issue for evaluating the work of judges in class action litigation. The court's interest may be at odds with the interests of the class in two ways. First, judges' overloaded dockets create an incentive to approve settlements to dispose of the complex litigation before them.[331] Second, judges who take an active role in crafting settlements may be influenced by their own contribution in overseeing the fairness hearing.[332] These issues have generally been glossed over in judicial review of settlements, but

---

powerful sanctions will actually be applied).

327. *See* Koniak & Cohen, *supra* note 9, at 1103. It is not clear whether the use of such strong deterrents against objectors is wise given the hurdles that objectors must already overcome. *See supra* note 100 (discussing heavy sanctions against objectors).

328. *See* Koniak & Cohen, *supra* note 9, at 1106-09.

329. *See* Lester Brickman, *Contingency Fee Abuses, Ethical Mandates, and the Disciplinary System: The Case Against Case-By-Case Enforcement*, 53 WASH. & LEE L. REV. 1339 (1996) (providing survey evidence to support claim that case-by-case disciplinary enforcement of contingency fee ethics rules does not work and arguing for systematic reform).

330. At the same time, we should not forget that class actions themselves are intended to serve a valuable compensation and deterrent function and that overly harsh penalties can result in an inefficient chilling of one of the only avenues for consumer redress of wrongs. Because we as a society have decided to address most consumer (as well as tort and other) wrongs through the mechanism of group litigation, the elimination or limitation on group litigation in the absence of legislatively imposed penalties or regulatory schemes is likely to result in no enforcement or deterrence of these individually small but collectively substantial wrongs.

331. *See* Koniak & Cohen, *supra* note 318, at 151 (discussing the possibility that judges seek to approve class action settlements in order to clear their dockets).

332. As Minow points out, judges both declare the law and resolve disputes, and these roles are not always coextensive. Where judges are in their role of resolving disputes, "confidence in the judge becomes the central point of contention." Minow, *supra* note 217, at 2027. This is because the "settlement process permits room for personal persuasion, input, or pressure by the judge." *Id.* at 2028.

should not be.

Expertise and independence are linked to the extent that the most experienced judges are also those most deeply involved in the structuring of settlements. Such judges may have difficulty stepping back at the 23(e) hearing and making an independent assessment of the settlement. Although critics have expressed concern over judges overstepping their bounds, preventing judges from being involved in settlement structuring would be too great a detriment to the governance of class actions. Class members would likely be worse off without judicial expertise in dispute resolution, especially with respect to attorneys' fees questions in which there is a structural incentive for class counsel to engage in self-dealing. The newly adopted changes to Rule 23, now requiring fee motions in class action cases, signal the rule-makers' intentions to expand judicial involvement and oversight.[333]

One method of ensuring judicial independence is to separate the judicial function of overseeing litigation from that of settlement approval. Under such a regime, the judge who oversees the negotiation, and is likely to try the case, would not be the same judge conducting the 23(e) hearing.[334] The 23(e) judge would be less subject to the influences of either the prospect of having to try a complex case or her own intervention in resolving the dispute. The appeal of this solution would depend in part on the cost of gaining case-specific knowledge.

Another method is to require that all settlement negotiations be handled by a magistrate judge, rather than making use of a magistrate discretionary. This solution would have limited effect because it would not remove the incentive for the Rule 23(e) judge to remove the case from his docket by approving settlement. Requiring a separate hearing before a new 23(e) judge would counteract the role of the judge in developing settlements and remove the judicial incentive to dispense with complex cases through inadequate settlements. At the same time, such a solution would permit the same judge who influenced the dynamics of settlement, in part because of the parties' predictions of what the judge would do at trial, to continue involvement in the case if settlement is not approved.[335]

There are significant problems with the idea of a Rule 23(e) judge, however. First, to adopt such a scheme is to admit that judges are driven by incentives and influenced by their own deep involvement in cases. It may be perceived as an affront to the ideal of judicial objectivity and independence. Second, a two-judge solution may also be inefficient because of the costs and time for the 23(e) judge to familiarize herself with a complex case. It is difficult to tell whether the costs

---

333. *See* FED. R. CIV. P. 23(h) discussed *supra* note 271.

334. Minow proposed a similar experiment, asking "[w]ould it be possible to ready a different judge to proceed with trial so the judge presiding over the settlement process could distinguish that role from the task of judging?" Minow, *supra* note 217, at 2029. This is different from the current practice of permitting a Multi-District Litigation judge to oversee a case until trial in that it separates the judges' negotiation and settlement approval functions.

335. This would resolve a problem, pointed to by Minow, that a two judge solution may prove "destructive to the dynamics of settlement (if it depends in part on predictions of what actually would happen at trial)." *Id.*

of familiarizing a 23(e) judge would exceed the benefits of more objective and independent review of settlements and the freedom it would give the presiding judge to influence settlement discussions.

This tension between judicial dispute resolution and judicial independence is not resolved by the availability of an appeal. The deferential standard for reviewing settlement approval prevents a critical review of settlements that might compel judges to exercise more care *ex ante* to reject unfair settlements.[336] There are two justifications courts provide for this deferential standard of review: first, that the trial judge is most familiar with the litigants, their strategies, positions and proofs, and second, the public policy presumption in favor of settlements.[337] The deferential standard may also be justified by the analogy of the class action to a "quasi-administrative proceeding."[338]

The first justification flips the criticism articulated above. Instead of seeing familiarity with the case as a form of bias, the prevailing doctrine of judicial review of class actions views familiarity purely from an efficiency perspective. Even taken at face value, however, this justification loses force where the judge below does not take the fairness hearing seriously.[339] Active involvement and familiarity with the case come at the cost of impartiality and independence. The second justification, a presumption in favor of settlements, requires a logical leap in the class action context. The intuition behind the law favoring settlements must be a contractual one, that private ordering is more efficient, saves judicial and public resources, and resolves conflicts to the satisfaction of both contracting parties.[340] But in the class action context, because of the inherent agent-principal problems, the settlement does not represent a contract between the class and defendant but between the *class counsel* and defendant. For this reason, courts

---

336. "[E]valuation of [a] proposed settlement in this type of litigation . . . requires an amalgam of delicate balancing, gross approximations and rough justice," and the trial court's ruling on the adequacy of a proposed compromise is given great deference. Detroit v. Grinnell Corp., 356 F. Supp. 1380, 1385 (S.D.N.Y. 1972), *aff'd in part and rev'd in part*, 495 F.2d 448 (2d Cir. 1974).

337. *See, e.g.*, Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 630 (1997) (Breyer, J., dissenting in part) ("The law gives broad leeway to district courts in making class certification decisions, and their judgments are to be reviewed by the court of appeals only for abuse of discretion."); Califano v. Yamasaki, 442 U.S. 682 (1979) (applying abuse of discretion standard); Petrovic v. Amoco Oil Co., 200 F.3d 1140, 1148 (8th Cir. 1999) (applying abuse of discretion standard and citing Grunin v. Int'l House of Pancakes, 513 F.2d 114, 123 (8th Cir.), *cert. denied*, 423 U.S. 864 (1975)); Little Rock Sch. Dist. v. Pulaski County Special Sch. Dist. No. 1, 921 F.2d 1371, 1388 (8th Cir. 1990); *In re* Equity Funding Corp. of Am. Sec. Litig., 603 F.2d 1353, 1362 (9th Cir. 1979) ("A district court's approval of a class action settlement and plan of allocation of proceeds among plaintiff classes is to be reviewed under the abuse of discretion standard.").

338. *See* Phillips Petroleum v. Shutts, 472 U.S. 797, 808 (1985) (describing the class action as a "quasi-administrative proceeding" from class members' point of view); Richard Nagareda, *Turning from Tort to Administration*, 94 MICH. L. REV. 899 (1996) (advocating application of the "hard look" doctrine to mass tort class actions).

339. *See supra* note 13 (discussing limitations of fairness hearings).

340. *But see* Owen M. Fiss, *Against Settlement*, 93 YALE L. J. 1073 (1984).

have held that district court judges have a fiduciary duty to the class and courts have developed a number of procedural and substantive factors to review prior to approving a settlement.[341]

An alternative is a *de novo* standard of review for class action settlement approval. Under this proposal, the problems surrounding class actions are viewed as so intractable and damaging to the perception of our justice system that a higher standard of review is warranted. It seems likely, however, that the same disincentives to reject settlements that apply to district court judges will be applicable at the appellate level as well. Appellate judges are no less likely to want to reduce their docket, especially if they are inundated with class action settlement appeals. Nor are appellate judges going to be less reticent than district court judges to re-open a complex matter that has been resolved.

A third option may be an intermediate standard of review.[342] This approach would vary the deference of the appellate courts' standard of review depending on the strength of the protections afforded the class. It would require *de novo* review of the process and protections afforded class members. The measure by which these protections would be judged should be the governance principles articulated in this section—disclosure, adversarial process, expertise and independent leadership—and correlating mechanisms to fulfill each of them. Thereafter, appellate courts would review *de novo* the district court's substantive fairness determinations in cases where governance is otherwise weak. Examples include cases without objectors, where the fee award was protected from an adversarial process because of a 'clear sailing' provision in the settlement agreement,[343] or where no separate Rule 23(e) judge reviewed the settlement prior to approval. Especially in complex and difficult cases, the standard of review should reflect the level of protection received by the class members. The better the mechanisms protecting the class in the first instance, the more deferential the standard of review should be. The underlying assumption of this proposal is that due process requires that the four principles of governance be met robustly in order to bind absent class members.[344]

---

341. *See, e.g.*, City of Detroit v. Grinnell Corp., 495 F.2d 448, 463 (2d Cir. 1974) (listing nine substantive factors and three procedural factors).

342. The dissent in *Amchem* indicated that the majority was indeed applying an intermediate standard of review and that is the reason why the court found the settlement to be patently unfair. *See* Amchem Prods., Inc. v. Windsor, 521 U.S. 591, 630-31 (1997) (Breyer, J., dissenting) ("These difficulties flow from the majority's review of what are highly fact-based, complex, and difficult matters, matters that are inappropriate for initial review before this Court."). *Id.* at 630.

343. A "clear sailing" provision is one in which the defendant promises not to contest class counsel's fee request. Such a provision, because it prohibits defendant from objecting to exorbitant fees, exacerbates the conflict of interest between class counsel and the class and further erodes the adversarial aspects of the fairness hearing.

344. This would mean that the right to collateral attack would depend on whether the four principles of governance were adequately reviewed, because collateral attack on a judgment binding class members is available where there has been a violation of due process. For enlightening discussions on this topic, *see* Monaghan, *supra* note 6 and Kahan & Silberman, *supra* note 47.

Like other actors in the class context, the independence of objectors and other intervenors from external influence—by defendants or class counsel—should also be measured. While objectors should be encouraged, they must also be monitored and subjected to the same governance requirements as class counsel. This is because the very leverage that permits objectors to destabilize settlements and obtain gains for the class also allows them to draw scarce resources from the class members with the threat of holding hard-fought settlements hostage. The independence of community groups and states' attorneys general may be at issue, particularly in settlements that provide for *cy pres* funding for those groups or give away benefits to persons outside the class. State attorneys general or community representatives, for example, may be inclined to support settlements that give away money to their constituencies even if they do not compensate the consumers harmed.[345] Independence of *guardians ad litem* may be questioned where they are appointed by the defendants and class counsel, or where they are paid out of the settlement fund or only paid if the settlement is approved. Courts should scrutinize the interests of intervenors in settlements that do not provide cash payments to claimants particularly carefully. Side agreements between objectors and class counsel at any stage of the litigation should be reviewed with close attention by courts.[346]

Control of objectors' counsel's fees is another means to assure independence.[347] Objectors' counsel have a perverse incentive to get a settlement approved (albeit with their proposed changes) so that they can get paid. If they succeed in scuttling a settlement, they will ordinarily not be paid unless they manage to strike a new deal with the defendants. Thus the least mercenary objectors, such as public interest organizations, are least likely to be paid for their efforts. Some might argue that this incentive is not perverse but in fact beneficial, because it pushes objectors to try to work within a hard-fought settlement and save the costs of renegotiation. Nevertheless, this incentive is just as likely to allow defendants and class counsel to get away with too much. This problem might be solved by requiring defendants and class counsel to post a bond. If the court rejects the settlement as unfair to class members, objectors' counsel who succeeded would be paid a reasonable amount out of that bond. This would put some incentive in the hands of defendants not to go forward with settlements that they know to be patently unfair.[348]

---

345. *See, e.g.*, *In re* Toys "R" Us Antitrust Litig., 191 F.R.D. 347 (E.D.N.Y. 2000) (attorneys general supported settlement provided for contribution of toys worth $36.6 million to states).

346. This would require a departure from the current practice and the revised rules, which currently only require the reporting of the existence of such side-settlements. *See* FED. R. CIV. P. 23(2) (2).

347. This criticism applies equally well to *guardians ad litem*.

348. Without a doubt, the defense bar would oppose such a requirement. The burden of the bond could be put on class counsel, but this would simply reinforce the antagonism between objectors' and plaintiffs' counsel and succeed in erasing the role of the defendant in forming the settlement. Policy makers should not forget that defendants also have a role in producing inadequate settlements.

A truly adversarial process is impossible without stringent court oversight over the independence of decisionmakers and intermediaries from self-interest and influence. This will depend in some part on the court's discretion, but the inquiry into the interests of these players in conjunction with the institution of specific procedural safeguards should eliminate some of the "chancellor's foot" problems created by unfettered judicial discretion.

CONCLUSION

The problems of class action abuses are a result of a failure of the existing governance regime. In place of that flawed regime, this Article proposes a system of governance for small claims damages class actions in the settlement process based on the four fundamental principles of disclosure, an actively adversarial process, agent expertise, and independence of decision-making. These principles form a comprehensive approach to class action settlement governance, requiring a number of mechanisms that should be implemented in all small claims damages class action settlements. By thinking of the class as a group in need of governance, and taking a comprehensive rather than piecemeal approach to determining what that governance should be, we can begin to have an understanding of the nature of the class as an organization, to improve the functioning of that organization, and to create a fair process more likely to yield fair substantive results.

ENRON, EPISTEMOLOGY, AND ACCOUNTABILITY: REGULATING IN A GLOBAL ECONOMY

ERICA BEECHER-MONAS*

TABLE OF CONTENTS

INTRODUCTION

Enron, WorldCom, Global Crossing . . . if the consequences were not despoiling the economy (and our retirement funds), the avalanche of corporate

---

* Associate Professor of Law, University of Arkansas at Little Rock William H. Bowen School of Law. J.S.D., Columbia School of Law; J.D., University of Miami School of Law. For comments on this Article, I am most grateful to Joan Heminway, Marco Janssen, and Mark Seidenfeld, as well as Martha Fineman and participants in the Feminist Legal Theory Workshop and Owen Jones and participants in the Fifth Annual Meeting of the Society for Evolutionary Analysis in Law. For her able and enthusiastic research assistance, I thank Dawn Peoples.

defections would be downright entertaining. Among the largest and most widely followed publicly held corporations, these kinds of firms were once thought the least likely to undertake risky actions that would lead to financial fraud. Nonetheless, large publicly held corporations increasingly have misrepresented their financial health.[1] As a result, corporate governance has become a key issue for reinvigorating investor confidence, impelling legislation, commentary and debate. This Article draws on insights from evolutionary biology, game theory, and cognitive decision theory to examine the current global crisis in corporate governance and proposes solutions to this predicament.

The current crisis in corporate governance is a consequence of a pervasive undermining of safeguards designed to prevent financial fraud. Congress, the Securities and Exchange Commission (SEC), and the courts have whittled away investor protections under the federal securities regulations using a combination of regulatory reforms and enactment of legislative and judicial barriers to enforcement mechanisms.[2] Despite the acknowledged importance of information to the functioning of efficient markets, the SEC has increasingly deregulated disclosure over the last two decades, partly for political reasons, and partly as an accommodation to globalization.[3] In addition, private litigation barriers have exacerbated the problem.[4] Commentators have justified deregulation of the

---

1. U.S. GEN. ACCOUNTING OFF., PUB. NO. GAO-03-138, REGULATORY RESPONSES AND REMAINING CHALLENGES 4 [hereinafter GAO Report]. The GAO Report found that between January 1997 and June 2002, 10% of all listed companies announced at least one financial statement restatement. *Id.* The Report finds a significant growth in fraudulent financial misrepresentations (showing 165% growth in financial statement restatements due to prior misrepresentations). *Id.* at 17. During this time period, the size of the typical restating company rose from an average (median) of $500 million ($143 million) in 1997 to $2 billion ($351 million) in 2002. *Id.* Issues involving revenue recognition accounted for nearly 38% of these restatements. *Id.* at 5.

2. *See, e.g.*, Larry E. Ribstein, *Market vs. Regulatory Responses to Corporate Fraud: A Critique of the Sarbanes-Oxley Act of 2002*, 28 J. CORP. L. 1, 17 (2002) (noting that "[r]ecent corporate frauds occurred following 1990s laws scaling back potential liability for corporate fraud"); John C. Coffee, Jr., *Understanding Enron: "It's About the Gatekeepers, Stupid,"* 57 BUS. LAW. 1403, 1409 (2002) (noting the deregulatory movement during the 1990s that sought to dismantle arguably obsolete regulatory provisions).

3. For example, Stephen Bainbridge, who argues in favor of deregulation, points to private securities litigation as a reason that mandatory disclosure is unnecessary. Stephen M. Bainbridge, *Mandatory Disclosure: A Behavioral Analysis*, 68 U. CIN. L. REV. 1023, 1033 (2000) (arguing for deregulation on the basis of the status quo bias of behavioral economics). What Bainbridge fails to address, however, are the severe impediments to private antifraud enforcement posed by the Private Securities Litigation Reform Act of 1995 (PSLRA).

4. The PSLRA (with its increased pleading standards) combined with judicial hostility to plaintiffs' civil actions, which make it more difficult for shareholders to remedy and deter nondisclosure, place huge obstacles against anti-fraud litigation and dismantle necessary safeguards to prevent corporate overreaching. *See* Douglas M. Branson, *Running the Gauntlet: A Description of the Arduous, and Now Often Fatal, Journey for Plaintiffs in Federal Securities Law Actions*, 65

securities industry using economics and game theoretic arguments. Their theory is that the market will force the evolution of efficient norms.[5] However, a proper game theoretic analysis suggests that the defections[6] of Enron, WorldCom and their ilk are predictable outcomes from deregulation. Further, Congress's proposed solution is unlikely to fix the problem. Both game theory and the complementary insights of cognitive psychology suggest that the corporate governance measures dictated by the Sarbanes-Oxley Act are an unrealistic and ineffective answer to the current financial scandals. Tracing the deregulatory impact on both public and private enforcement mechanisms, this Article argues that even the most efficient markets need strong investor protections,[7] and

---

U. CIN. L. REV. 3, 11 (1996). Some scholars contend that PSLRA has not decreased the level of meritorious filings, but their results are inconclusive. *See* Michael A. Perino, *Did the Private Securities Litigation Reform Act Work?*, 2003 U. ILL. L. REV. (forthcoming 2003) (noting the many variables that may affect the number of filings, and noting that "while it is difficult to assess the claim that there is more fraud now than there was prior to the PSLRA, the other explanations for the apparent increase in filings appear to be inadequate"); Ribstein, *supra* note 2, at 17 (arguing that "reduced liability risk may have encouraged fraudulent or shirking behavior in marginal situations where defrauding insiders or lax auditors had persuaded themselves that the likelihood of detection was low . . . [which] argues for reversing some aspects of PSLRA"). PSLRA was followed by the Securities Litigation Uniform Standards Act of 1998 (SLUSA), 15 U.S.C. § 77k(b) (2000). *See, e.g.*, Lander v. Hartford Life & Annuity Ins. Co., 251 F.3d 101 (2d Cir. 2001) (holding that SLUSA applied to a class action alleging misrepresentations in the sale of annuity contracts, which activated removal); *In re* BankAmerica Corp. Sec. Litig., 263 F.3d 795, 799-801 (8th Cir. 2001) (explaining that Congress was funneling class-action securities litigation into the federal courts). The Supreme Court has also had a part in diminishing enforcement, through its decision in *Central Bank of Denver v. First Interstate Bank of Denver*, 511 U.S. 164, 191 (1994) (holding that private fraud actions under § 10(b) and Rule 10b-5 cannot be brought under an aiding and abetting theory).

5. *See, e.g.*, Bainbridge, *supra* note 3, at 1033 (arguing from a game theoretic perspective that the market will force the evolution of efficient norms, making regulation unnecessary); Frank H. Easterbrook & Daniel R. Fischel, *Mandatory Disclosure and the Protection of Investors*, 70 VA. L. REV. 669, 682 (1984) (arguing from a law and economics perspective for deregulation); *see also* Bernard Black & Reinier Kraakman, *A Self-Enforcing Model of Corporate Law*, 109 HARV. L. REV. 1911, 1939-40 (1996) (contending that under conditions of efficiency a nonlegal sanctioning system will be sufficient).

6. Defection is a term used by game theorists to express a self-interested strategy that is strictly dominant in prisoner's dilemma games, in that it is the best choice for a player given every possible choice by another player. *See* DOUGLAS C. BAIRD ET AL., GAME THEORY AND THE LAW 36 (1994) (a "strictly dominant" strategy is the best choice for a player given every possible move by another player). In the context of issuer/investor interactions, the self-interested moves (defections) consist of management self-dealing and director passivity and nondisclosure.

7. *See generally* JOHN MCMILLAN, A NATURAL HISTORY OF MARKETS (2002) (arguing that for markets to thrive in a socially productive manner, they require constant government tinkering; and explaining that "the efficacy of the stock market varies with how activist the government is in setting the platform" since "countries with stronger investor protections have bigger capital markets").

contends that the resulting vacuum created a climate ripe for corporate malfeasance.

In response to these corporate debacles (and to Enron in particular), Congress passed the Sarbanes-Oxley Act,[8] and directed the Securities and Exchange Commission (SEC) to engage in rulemaking to address the perceived problems.[9] Among other changes, the Act requires increased independence of auditors, directors, and analysts; beefing up the disclosures required in annual reports; and changing accounting rules that permit special purpose entities to disguise losses. Congress's principal solution regarding corporate governance was to place the firm's audit committee in charge of the relationship between the firm and its auditors.[10] In addition, the audit committee must monitor a system of internal accounting controls—put in place by the chief executive and chief financial officers—to ensure that the flow of information reaches them.[11] Each annual report must contain an internal control report.[12] Thus, the Sarbanes-Oxley Act makes the audit committee responsible for corporate financial disclosures.[13]

This legislative response is unlikely to accomplish the necessary change. Although financial information, current business developments, and future plans are foundational information for investor decisionmaking,[14] the dynamics of

---

8. *See* Michael Schroeder, *The Economy: SEC Orders New Disclosures on Company Earnings*, WALL ST. J., Jan. 16, 2003, at A2 ("Responding to recent corporate scandals . . . federal securities regulators ordered new disclosure rules to clamp down on an accounting practice that companies have increasingly used to paint rosy financial results. . . . The changes were ordered by Congress under the S-O Act, a sweeping corporate accounting-overhaul law . . . .").

9. Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, 116 Stat. 745; *cf. Enron Fallout: Public Policy Consequences of Enron's Collapse*, 34 Sec. Reg. & L. Rep. S-5 (BNA) (Mar. 4, 2002) (discussing proposed changes).

10. Sarbanes-Oxley Act § 301. The auditor is to be hired by and report directly to the audit committee, which must be composed of independent directors, at least one of whom must be a financial expert. Sarbanes-Oxley Act § 407 (noting that if the audit committee has no financial expert it must disclose the reasons for the absence).

11. The chief executive officer and chief financial officer must set up a compliance system, and certify that they have disclosed any deficiencies, fraud, or significant changes in the internal controls to the auditors and the audit committee. Sarbanes-Oxley Act § 302(a)(4)-(6). The chief executive officer and chief financial officer must certify in each annual and quarterly report that they have reviewed the report, that it is true (to their knowledge), that the financial statements and other financial information fairly present the financial condition of the company, that they have established and maintain internal controls designed to ensure that material information is made known to them (and any deficiencies have been disclosed to the auditor and the audit committee). Sarbanes-Oxley Act § 302.

12. Sarbanes-Oxley Act § 404.

13. *See* Sarbanes-Oxley 2002 § 301(4)(A) (the credit committee must establish procedures for resolving complaints about financial matters); § 301(2) (the audit committee must resolve disagreements between auditors and management); § 302 (CEO and CFO must report to audit committee deficiencies, fraud, or significant changes in internal control system).

14. *See* STEPHEN M. BAINBRIDGE, CORPORATION LAW AND ECONOMICS 300 (2002)

small group interactions illuminate the problematic aspects of the Sarbanes-Oxley Act's attempts to achieve accountability through compliance programs and independent directors. Directors undertake decisions and actions as a group.[15] This decisionmaking context has important consequences for any attempt to resolve the agency problems arising from the separation of ownership and control.[16] Group decisions, while offering many advantages over individual decisions, have limitations as well. Simply putting independent directors in charge of monitoring the corporation will not solve the problems inherent in group decisionmaking, as evolutionary game theory and cognitive psychology demonstrate.

This Article proceeds in five parts. Following the Introduction, Part I outlines the theoretical basis for a mandatory disclosure regime in securities law, the bureaucratic problem of ensuring that those who are nominally in charge of corporate decisions have access to the kinds of information they need, and the congressional solution of placing the audit committee in charge of corporate compliance. Part II discusses evolutionary game theory and the importance of regulatory structure for optimal social gains to occur. Part III discusses the Sarbanes-Oxley Act's solution of internal control system disclosure and explores the dynamics of organizational behavior in the context of financial reporting decisions under conditions of financial stress. Drawing on evolutionary game theory and cognitive psychology, Part IV proposes a self-insurance solution for large publicly held corporations that would cover independent directors' liability for financial misrepresentations that involved recklessness (but not self-dealing). This Article concludes that undermining enforcement mechanisms and decreasing disclosure obligations may have the kinds of adverse consequences the Enron implosion exemplifies. Law has an important function not only in solving information asymmetries, but also in altering players' incentives to make socially valuable transactions more likely and in channeling behavior. Both regulation and private enforcement are important components of efficient capital markets.

(observing that no one seriously disputes the importance of information as a solution to the basic agency problems arising from the separation of ownership and control); Ian Ayres & Eric Talley, *Solomonic Bargaining: Dividing a Legal Entitlement to Facilitate Coasean Trade*, 104 YALE L.J. 1027, 1039 (1995) (discussing the disclosure increasing effect of liability rules).

15. For example, the Delaware Code provides that the "vote of the majority of the directors present at a meeting at which a quorum is present shall be the act of the board of directors unless the certificate of incorporation or the bylaws shall require a vote of a greater number." DEL. CODE ANN. tit. 8, § 141(b) (2002).

16. The agency problem is a result, as Bearle and Means explained, of the separation of management and control. ADOLF A. BEARLE & GARDNER C. MEANS, THE MODERN CORPORATION AND PRIVATE PROPERTY 6 (1932) ("The separation of ownership from control produces a condition where the interests of owner and of ultimate manager may, and often do, diverge. . . .").

I. THE PROBLEM CONGRESS TRIED TO SOLVE

A. Divergent Incentives

In his article, *The Nature of the Firm*, Ronald Coase explained that a firm substitutes bureaucracy, hierarchy, and fiat for contract as a method of reducing transaction costs.[17] Businesses coordinate individuals who specialize, and whose activities relate to each other.[18] The problem is that the interests of these individuals may not always be aligned, creating agency costs. The arguments for imposing duties on corporate managers (officers and directors) on behalf of the shareholders are articulated either as a principal/agent relationship—which is problematic, because the director/shareholder relationship lacks most of the attributes of such a relationship[19]—or as a "nexus of contracts," in which the other firm participants demand contracts with the firm for payment before any payment can be made to the shareholders.[20]

Under the nexus of contracts theory, shareholders get to elect directors and impose fiduciary duties as an implied contractual exchange for accepting higher risk.[21] Under either concept, the interests of the residual claimants (the

---

17. Ronald H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386, 386-405 (1937).

18. Thus, the corporation is said to be a nexus of contracts, with its predominant feature being the separation of ownership and control. FRANK H. EASTERBROOK & DANIEL R. FISCHEL, THE ECONOMIC STRUCTURE OF CORPORATE LAW 1 (1999). The shareholders, who own the firm, and the managers, who run it, have divergent interests. Melvin A. Eisenberg, *The Structure of Corporation Law*, 89 COLUM. L. REV. 1461, 1471 (1989).

19. Margaret M. Blair & Lynn A. Stout, *Director Accountability and the Mediating Role of the Corporate Board*, 79 WASH. U. L.Q. 403, 409-11 (2001) (describing the fallacy of seeing the relationship as a principal/agent relationship). For example, shareholders, unlike principals, have no power to initiate corporate action, their vote is limited to choosing directors and to extraordinary board actions. *Id.*

20. *See* PAUL MILGROM & JOHN ROBERTS, ECONOMICS, ORGANIZATION & MANAGEMENT 20 (1992) (discussing the theory of the firm as a "nexus of contracts, treaties, and understandings among the individual members of the organization"); *see also* Blair & Stout, *supra* note 19, at 410 (describing shareholder primacy under economic theories as a consequence of seeing the relationship as a principal/agent relationship, but contending that the better theory of the firm is as a nexus of contracts, in which "nonshareholder participants in the firm (including bondholders, managers, and employees) demand contracts that require them to be fully compensated out of any revenues earned by the enterprise before any payments can be made to shareholders").

21. *See* EASTERBROOK & FISCHEL, *supra* note 18 (describing the theory of the firm as a nexus of contracts with shareholders as residual claimants). This is the shareholder primacy model of corporate structure. *See* BAINBRIDGE, *supra* note 14, at 10, 29. The director primacy model, on the other hand, sees the corporate bureaucracy as dominated by professional managers, with the directors acting as mere figureheads, and the shareholders as being largely irrelevant. Under this model, the directors are only accountable for increasing shareholder wealth. Whether in fact this is what happens, the laws of every state place a monitoring function on the board and fiduciary duties that run to the shareholders.

shareholders), the centralized management, and the monitors may diverge and monitors may shirk or self-deal (the classic problem of agency costs).[22] The corporate structure contemplates decisionmaking by professional manager's fiat, monitored by the board, in which the board acts and the shareholders react.[23] If an entity chooses the corporate form, all states place directors at the apex of the decisionmaking structure.[24] In other words, the board has the power to monitor and discipline management, to make policy, and to demand access to resources, such as legal and accounting advice.[25] Shareholders have the power to withdraw, vote, and enforce duties owed to them.[26]

Acknowledging that shareholders have such rights does not necessarily imply a regime of director liability. One of the common arguments against a shareholder primacy model is that because shareholders can reduce their risk by diversifying, there is no need for a rule of director liability.[27] This view contends that all directors should only be responsible for maximizing shareholder wealth, and if they fail to do so, the shareholders should sell. There are three responses to this argument. First, as Stephen Bainbridge explains, management misconduct is not a diversifiable risk.[28] Risk is defined by reference to the variance on return.[29] Misconduct does not affect variance, it erodes expected return.[30]

Second, a steady flow of truthful information into the primary and secondary

---

22. *See* BEARLE & MEANS, *supra* note 16, at 6.

23. The Delaware Code places the nexus of contracts squarely on the directors, requiring that the corporation's "business and affairs . . . shall be managed by or under the direction of a board of directors. . . ." DEL. CODE ANN. tit. 8, § 141(a) (2002). For a discussion of the nexus theory, see BAINBRIDGE, *supra* note 14, at 197-204. The concept of a monitoring board, although accepted by state corporation statutes, is controversial in practice. *See* Stephen M. Bainbridge, *Independent Directors and the ALI Corporate Governance Project*, 61 GEO. WASH. L. REV. 1034, 1048 (1993). For example, when Melvin Eisenberg attempted to create the concept of a monitoring board in the ALI Principles of Corporate Governance project, it created such a storm of controversy that he was forced to drop the word "monitoring" from the ALI. *Id.* at 1048.

24. This is probably not the reality of the situation. Most decisions undoubtedly are delegated except for extraordinary decisions. As the Enron directors explained to the congressional investigators, they had what was essentially a part-time job. GAO Report, *supra* note 1, at 17.

25. *See, e.g.*, MILGROM & ROBERTS, *supra* note 20, at 314 (noting that "[i]f any group could be considered to have residual control in a corporation, perhaps it might be the board of directors . . . [who] have the power to set dividends; to hire, fire, and set the compensation of the senior executives; to decide to enter new lines of business; to reject merger offers or instead approve and submit them to the stockholders; and so on").

26. *See id.* at 508 (discussing the shareholder options of exit and voice).

27. *See* BAINBRIDGE, *supra* note 14, at 263 (articulating the portfolio rationale for the business judgment rule).

28. *See id.* at 263 n.31.

29. *See* MILGROM & ROBERTS, *supra* note 20, at 461 ("Risk is measured by the variance of the investment returns or their standard deviation (the square root of the variance).").

30. *See* BAINBRIDGE *supra* note 14, at 263 n.31.

markets helps these markets grow.[31] If the market can effectively price, it can effectively allocate capital investment.[32] Capital will flow in the direction indicated by prices.[33] If there is undisclosed management fraud, investors and society suffer. Thus, sanctions are an important way of deterring misconduct. In addition, because it is common knowledge that directors and managers have incentives to withhold bad news, law gives them a way of assuring the market that they will not withhold it. Moreover, by telling issuers what information must be disclosed, regulation serves a channeling function.

Only the strongest form of the efficient market hypothesis suggests that stock prices reflect the securities' intrinsic value.[34] Even voluntary disclosure advocates acknowledge the importance of disclosure.[35] Withholding information, or providing incorrect information, is not an option for efficient markets. Voluntary disclosure theorists simply argue that market forces will provide sufficient incentives for firms to provide optimal levels of information.[36] This argument, however, assumes that capital financing endeavors are infinitely repeated games in which players can verify information and punish firms that provide inaccurate or too little information.[37] Game theory explains, however,

---

31. If accurate information is not made rapidly available, the markets flounder. *See, e.g.*, Marc I. Steinberg, *Curtailing Investor Protection Under the Securities Laws: Good for the Economy?*, 55 SMU L. REV. 347, 347 (2002) (noting that "[r]elatively efficient trading markets are based on a disclosure regime where transactions are expeditiously executed and competitively priced").

32. Empirical studies of stock prices before and after the announcement of a significant event show not only that stock prices respond, but that they do so rapidly (as they did, for example, following Enron's announcement that its earnings would have to be restated for the past four years).

33. The price of stock, according to the efficient market hypothesis, reflects a consensus of market participants about the present value of a future income stream. Although the efficient market hypothesis supports the rapid incorporation of publicly available information into the stock prices, this presupposes that information is made available. *See, e.g.*, MILGROM & ROBERTS, *supra* note 20, at 467-69 (explaining that the efficient market hypothesis, under the strong form—which says stock prices reflect all information—or the moderate form—which says that stock prices reflect all publicly available information—means that stock price is a proxy for managerial performance).

34. *See* Eugene Fama, *Efficient Capital Markets II*, 46 J. FIN. 1575 (1991). The strong form of the efficient market hypothesis has been challenged by empirical studies showing that "variations in stock prices were much too large to be explained as responses to changing expectations about future dividends." MILGROM & ROBERTS, *supra* note 20, at 470 (citing studies). Empirical support for the weak form of efficiency—that publicly available information is incorporated rapidly into the price of stock—has support. *See id.* ("The most recent econometric studies tend to support the view that the Weak Form Efficient Market Hypothesis is not fully consistent with the evidence, but that the deviations from pricing efficiency are not so great as to contradict the hypothesis 'grossly.'").

35. *See* EASTERBROOK & FISCHEL, *supra* note 18, at 288-89.

36. *See* Black & Kraakman, *supra* note 5, at 1939-40 (contending that under conditions of efficiency a nonlegal sanctioning system will work).

37. *See* Mitu Gulati, *When Corporate Managers Fear a Good Thing is Coming to an End: The Case of Interim Nondisclosure*, 46 UCLA L. REV. 675, 691 (1999) (noting that voluntary

that verifying information is a classic problem. Even sophisticated shareholders may have difficulty discerning whether financial statements are inaccurate.[38]

Third, the inability of the market to verify on a timely basis when defection has occurred means that sanctions will not be able to deter defections. Moreover, market incentives that sanction companies for misbehavior do not necessarily affect managerial behavior. In theory, poor managerial decisions should be reflected, at least after the fact, in stock prices, placing managers who make poor decisions in danger of being replaced.[39] However, because the market cannot distinguish between the consequences of managerial decisions and forces outside of management control, monitoring price is not equivalent to monitoring managerial performance. It is difficult to distinguish or differentiate managerial performance from breach.[40] Bad decisions and good decisions with bad consequences are hard for the market to distinguish.[41] Moreover, it may be that the reason the price declined had more to do with information about the firm than about the decisional performance of its managers.[42]

In addition, noise theory suggests that although stock prices reflect

---

disclosure theory is based on "two central assumptions . . . that companies and the market are playing an infinitely repeated game in which the benefits of cheating once are far outweighed by the reputational costs or other non-legal sanctions the company will have to bear in later transactions . . . [and] that the market can verify when the company has cheated it"). The assumption that raising capital is an infinitely repeated game is somewhat problematic, since, as Lynn Stout pointed out, it is a rare occurrence for corporations to raise capital through new equity. *See* Lynn A. Stout, *The Unimportance of Being Efficient: An Economic Analysis of Stock Market Pricing and Securities Regulation*, 87 MICH. L. REV. 613, 644-51 (1988). Nevertheless, firms are repeat players in the market in the sense that even private funding sources will be monitoring stock price. *See* Gulati, *supra*, at 730 n.160 (observing that "stock price can probably affect secondary sources of funds such as the private debt market"). Indeed, Enron's ability to obtain financing was linked to its stock price and that was the trigger for its bankruptcy.

38. *See, e.g.*, William W. Bratton, *Enron and the Dark Side of Shareholder Value*, 76 TUL. L. REV. 1275, 1338 (2002) (noting that despite the opacity of Enron's financial statements, investment professionals failed to question them).

39. *See* Stephen A. Ross, *Disclosure Regulation in Financial Markets: Implications of Modern Finance Theory and Signaling Theory*, in ISSUES IN FINANCIAL REGULATION 177, 183-88 (Franklin R. Edwards, ed. 1979) (explaining ways in which markets monitor managers); Jeffrey N. Gordon, *The Shaping Force of Corporate Law in the New Economic Order*, 31 U. RICH. L. REV. 1473, 1486 (1997) (acknowledging that stock prices are "noisy" and "imperfect signals" in monitoring managers).

40. *See* Daniel R. Fischel, *Market Evidence in Corporate Law*, 69 U. CHI. L. REV. 941, 959 (2002) (arguing that although "an increase in stock price . . . should have defeated a claim for liability and damages" under corporate law, a decline in stock prices does not "establish or even suggest that the corporate managers who made the decision should be liable in damages").

41. *Id.*

42. *Id.* at 960 (contending that if the decisional "goal is otherwise lawful, and no other grounds exist . . . for attacking what managers did, the lower stock price should be irrelevant").

information, they do so with some over- and under-reaction.[43] As a result, prices are an imperfect surrogate for managerial behavior.[44] Share prices do not necessarily provide guidance in evaluating corporate decisions.[45]

Thus, assuming shareholders have the power to enforce their contractual rights, they need to have information about how well the directors are performing their oversight duties.[46] Disclosure is said to be the essence of the fiduciary

---

43. *See, e.g.*, Donald C. Langevoort, *Theories, Assumptions, and Securities Regulation: Market Efficiency Revisited*, 140 U. PA. L. REV. 851, 857-72 (1992); Robert Leroy, *Efficient Capital Markets and Martingales*, 27 J. ECON. LIT. 1583, 1612 (1989) (remarking that "by renaming irrational trading 'noise' trading, [Fischer Black, who coined the term] avoided the I-word, thereby sanitizing irrationality and rendering it palatable to many analysts"). Not only do many investors ignore the supposed efficiency of the markets in their investment behavior, commonly believing stocks to be mispriced, *see, e.g.*, Barber & Odean, *Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors*, 55 J. FIN. 773 (2000), but they also act as a herd, following the latest trends and rumors (as evident by the tech boom of the past five years and Greenspan's frequent tirades against the irrational exuberance of the markets). *See* JOHN M. KEYNES, THE GENERAL THEORY OF EMPLOYMENT, INTEREST AND MONEY 156 (1936) (describing Keynes' theory that stock prices reflect investors' herd behavior and strategic assessments of what the crowd would do). Although the EMH predicts that over- or under-reactions of pricing to information will be short lived, because arbitragers will take advantage of the mispricing, this turns out—empirically—not to be the case, because of the impossibility of predicting when the mispricing will cease. *See, e.g.*, Gulati, *supra* note 37 (noting "empirical evidence showing that financial markets both under- and overreact to information about firms") (citing studies). These observations indicate that prices do not move smoothly to some equilibrium point reflecting their intrinsic value. *See* Daniel et al., *Investor Psychology and Security Markets Under- and Overreacting*, 53 J. FIN. 1839 (1998). For example, a study of 66,000 accounts at a discount brokerage found that the most frequent traders received a return of 11.4%, the samples' average return was 16.4%, and the market return during the study period was 17.9%. *See* Barber & Odean, *Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors*, 55 J. FIN. 773 (2000).

44. *See* Louis Lowenstein, *Financial Transparency and Corporate Governance: You Manage What You Measure*, 96 COLUM. L. REV. 1335, 1343-44 (1996) (arguing that stock prices are an inadequate measure of performance); Andrei Schleiffer & Lawrence H. Summers, *The Noise Trader Approach to Finance*, 4 J. ECON. PERSP. 19, 19-33 (1990) (describing noise theory).

45. *See* Noel Gaston, *Efficiency Wages, Managerial Discretion, and the Fear of Bankruptcy*, 33 J. ECON. BEHAV. & ORG. 41, 42 (1997) (noting that a number of recent theoretical models are based on the "recognition that share prices do not necessarily provide perfect guidance in evaluating corporate decisions").

46. *See* Armen A. Alchian & Harold Demsetz, *Production, Information Costs, and Economic Organization*, 62 AM. ECON. REV. 777 (1972) (explaining that final monitoring authority is given to the residual claimants in order to encourage detection and punishment of shirking). The firm's nominal owners, the shareholders, have delegated control over daily operations and long-term policy to directors, who in turn delegate these powers to firm managers, because of the monitoring difficulty. *See* BAINBRIDGE, *supra* note 14, at 512 (discussing the impediments to shareholder democracy).

obligation,[47] and shareholders will need a mechanism other than the power to withdraw in order to enforce it. Disclosure, however, is not an unlimited obligation.[48] Rather, a duty to disclose arises only if a statute or regulation requires disclosure, or if a corporation makes incomplete or misleading disclosures.[49]

The securities laws place disclosure obligations on corporate managers and directors in the form of reporting obligations and liability under the antifraud provisions.[50] State law regulates disclosure through the duty of care and the duty of loyalty. Because the business judgment rule protects even bad decisions of the board, however, the duty of care means only that the directors must engage in a good faith decision process and disclose the process that they utilized.[51] In addition, directors who have a duty to disclose may not knowingly or deliberately fail to disclose facts that they know are material.[52] The duty of loyalty prohibits director (and manager) self-dealing, which means that for a board decision to be self-interested, a majority of the board must be materially affected by the decision in a way not shared by the firm or the shareholders.[53] Most importantly, for liability to ensue, the decision must not have been approved by the shareholders after full disclosure.[54]

---

47. Robert W. Hillman, *Business Partners as Fiduciaries: Reflections on the Limits of Doctrine*, 22 CARDOZO L. REV. 51, 69 (2000) (explaining that "disclosure is the essence of the duty of a fiduciary").

48. As the Supreme Court explained in the context of the duty to disclose or abstain from trading, mere possession of material nonpublic information does not trigger a duty to disclose. Chiarella v. United States, 445 U.S. 222, 235 (1980).

49. *Id.*

50. *See* Basic, Inc. v. Levinson, 485 U.S. 224, 239 n.17 (1988) ("Silence, absent a duty to disclose, is not misleading . . . .").

51. *See* BAINBRIDGE, *supra* note 14, at 297 (discussing the "sticking point . . . [as] the adequacy of disclosure" because "[i]t is hard to imagine a board disclosure along the lines of: 'we're very sorry but we violated the duty of care in the following particulars, which we now describe at great length'").

52. *See In re* BankAmerica Corp. Sec. Litig., 78 F. Supp. 2d 976, 992-93 (E.D. Mo. 1999) (discussing the necessity of a duty to disclose before triggering the exception for knowingly or deliberately failing to disclose material facts under Delaware Code provision section 102(b)(7) that permits corporations to exculpate directors for breaches of fiduciary duty).

53. Directors are deemed disinterested when they "neither appear on both sides of a transaction nor expect to derive any personal financial benefit from it in the sense of self-dealing, as opposed to a benefit which devolves upon the corporation or all stockholders generally." Aronson v. Lewis, 473 A.2d 805, 812 (Del. 1984) (citations omitted). *See also* Rales v. Blasband, 634 A.2d 927, 936 (Del. 1993) (stating "[d]irectorial interest also exists where a corporate decision will have a materially detrimental impact on a director, but not on the corporation and the stockholders").

54. *See* Smith v. Van Gorkom, 488 A.2d 858, 890 (Del. 1985) (stating that failure to disclose material information, including the "fact that the Board had no reasonably adequate information indicative of the intrinsic value of the Company" made shareholder ratification unavailing).

Whether the underlying misconduct was self-dealing or complete lack of oversight, the disclosure problem is nearly always going to be misrepresentation by the board about their decision process or action, or a failure to disclose the basis for their decision.[55] From these basic precepts, the duty of candor and the duty to disclose arise.[56] These duties acknowledge the widely divergent incentives with respect to disclosure between corporations and their managers and investors: investors uniformly prefer more; directors and managers would prefer to withhold adverse information or information that could affect the firm's competitive situation (and their own self-interests), while disclosing favorable information to attract investors.[57]

In order to trigger a disclosure duty, however, the directors must have information to disclose or be reckless in failing to obtain it.[58] The problem in large organizations is how to ensure a flow of information both to those who manage the daily operations of the firm and to those who invest in the firm. A further problem is how to make those managing the firm accountable for acting (or failing to act) on the information they obtain, without unduly undermining the discretionary authority firm managers need to make the firm profitable.[59] Ensuring that information reaches the directors is the basis for the Sarbanes-

---

55. As Professor Langevoort explains it, "by exposing the 'lemons' in the market basket via well-enforced disclosure requirements, it creates an environment in which both markets and the better issuers and managers can flourish." Donald C. Langevoort, *Seeking Sunlight in Santa Fe's Shadow: The SEC's Pursuit of Managerial Accountability*, 79 WASH. U. L.Q. 449, 489 (2001) (arguing that securities law and state corporate law have complementary aims in controlling agency costs in the public corporation).

56. RESTATEMENT (SECOND) OF AGENCY § 381 (1958) (stating that an agent has the duty to disclose all matters relating to the agency).

57. Amir N. Licht, *Games Commissions Play: 2x2 Games of International Securities Regulation*, 24 YALE J. INT'L L. 61, 85-86 (1999).

58. Hence the protestations of the Enron board members that they knew nothing about the dire straits of the corporation and that management withheld key information from them. *See* Report of the Senate Permanent Subcommittee on Investigations on the Role of the Board of Directors in Enron's Collapse, S. REP. NO. 107-70, at 9 (2002) [hereinafter Enron Report].

59. *See generally* Michael P. Dooley, *Two Models of Corporate Governance*, 47 BUS. LAW. 461 (1992) (explaining the central corporate governance question as achieving the proper mix of accountability and discretion). This is the reason for the business judgment rule in state corporation law. *See, e.g.*, Gagliardi v. TriFoods Int'l, 683 A.2d 1049, 1052 (Del. Ch. 1996) (explaining that because shareholders "shouldn't rationally want" directors to be risk averse, the courts will abstain from interfering with business decisions); Joy v. North, 692 F.2d 880, 885 (2d Cir. 1982), *cert. denied*, 460 U.S. 1051 (1983) (noting that "the business judgment rule merely recognizes a certain voluntariness in undertaking the risk of bad business decisions"). Even bad decisions are virtually unreviewable, absent fraud, self-dealing, and such utter abdication of oversight responsibilities as to amount to aiding and abetting. *See, e.g.*, Francis v. United Jersey Bank, 432 A.2d 814 (N.J. 1981) (holding duty of care requires directors to inform themselves about the affairs of the firm, so that a widow on notice from her deceased husband that her sons were liable to loot the company was under a duty to take action to prevent the loss).

Oxley Act's mandate that managers establish and maintain internal controls, and disclose their evaluation of the corporate compliance program.[60]

B. Financial Misrepresentation: What Enron Didn't Say

The special purpose entities that ultimately destroyed Enron are a common type of asset securitization, involving a transaction between a selling company and an entity created for the sole purpose of buying its assets.[61] In accounting terms, the asset moves off the selling company's books, and is frequently followed by a swap agreement, in which the seller reassumes risks tied to the asset.[62] Enron engaged in a number of these transactions, with a twist.[63] Its swap agreements tied buyback provisions to its stock value, and permitted Enron to retain abnormally high risks.[64] Instead of using these entities for "legitimate purposes of achieving asset-liability matching, lowering funding costs, or improving liquidity," Enron used these entities to achieve an accounting result lacking in economic substance.[65]

Hundreds of millions of dollars a year were kept off the Enron balance sheets by using these special purpose entities.[66] The Financial Accounting Standards Board (FASB) is trying to close such loopholes by requiring that off-balance sheet partnerships be consolidated in the parent company's books unless there is an investment by outside parties equaling 10% of the total capital (the former

---

60. Sarbanes-Oxley Act of 2002 § 302. Notably, reporting companies already had to implement accounting controls under the Foreign Corrupt Practices Act, 1934 Securities Exchange Act § 13(b)(2)(A) (which required issuers to keep records "which, in reasonable detail, accurately and fairly reflect the transactions and dispositions of the assets of the issuer") and the 1934 Securities Exchange Act § 13(b)(2)(B) (which required issuers to maintain a system of internal financial controls to assure that transactions are properly authorized and that issuers have reliable information). No scienter is required in order to prove a violation of these requirements. SEC v. World-Wide Coin Invs., Ltd., 567 F. Supp. 724, 749-50 (N.D. Ga. 1983). Rule 13b2-1, promulgated under § 13(b)(2), prohibits falsification of the books and records of reporting companies. *Id.* at 746. Moreover, internal controls are effectively required under the Federal Sentencing Guidelines, which permit the existence of an effective system to be a mitigating factor in sentencing. UNITED STATES SENTENCING GUIDELINES § 8C2.5(f) (2002) (reducing base fines by up to 60%).

61. Jenny B. Davis, *The Enron Factor*, 88 A.B.A. J. 40, 42 (2002).

62. *See id.* at 42 (arguing that although such transactions may "sound suspect," they are "valid financial constructions—so long as the seller company and the SPE are independent of each other and the SPE has assumed the risks a buyer would normally assume").

63. *See* Jonathan C. Lipson, *Enron, Asset Securitization and Bankruptcy Reform: Dead or Dormant?*, 11 J. BANKR. L. & PRAC. 101, 112 (2002) (discussing Enron's use of special purpose entities to manipulate accounting disclosures).

64. Davis, *supra* note 61, at 42.

65. Kip Betz, *Securitizations: Panelists Call SPEs Legitimate Tool; Only Small Portion Used Improperly*, SEC. L. DAILY, Mar. 31, 2002, at D7.

66. 34 Sec. Reg. & L. Rep. S-11 (BNA) (Mar. 4, 2002).

rules required only 3% outside investment).[67] Also in the works are new rules regarding guarantees and accompanying obligations, with a proposal that the guarantor stand ready to perform over the term of the guarantee in the event of adverse financial conditions and accompanying future payments.[68] If guarantees or other arrangements shield investors from losses, the structure would have to be consolidated.[69]

The International Accounting Standards Board (IASB) has much tougher rules for off-balance sheet reporting.[70] Had the United States adopted the IASB rules, Enron would have had to disclose its special purpose entities.[71] Even so, IASB Chair David Tweedie acknowledged that "there may be ways in which the [international] rules could be 'strengthened or clarified.'"[72]

Enron not only failed to disclose its financial instability, it also failed to reveal multiple layers of conflicts of interest.[73] For example, in order to improve

---

67. *Id.*

68. *Id.*

69. Steve Burkholder, *Accounting: FASB to Address Financial "Conduits" in SPE Effort; Work on Guarantees Continues*, BANKING DAILY, Mar. 17, 2002, at D6.

70. *See* Patrick Tracey, *Accounting: Global Standard Setters Shift Focus to Debts Left Off Balance Sheets*, Pension & Benefits Daily Rep. (BNA) (Feb. 21, 2002).

71. *See Questioning the Books: Enron Hoped to Sway Accounting Group*, WALL ST. J., Feb. 14, 2002, at A8 (opining that "[h]ad the U.S. adopted the IASB's stricter rules, Enron would have been required to disclose [its special purpose entities] in financial statements").

72. Tracey, *supra* note 70. Although the SEC is considering the move toward international accounting standards—recognizing that a single system promotes transparency by permitting investors to compare the financial statements of different companies—it nonetheless continues to insist on reconciliation with GAAP. SECURITIES & EXCHANGE COMMISSION, NO. 7801, CONCEPT RELEASE: REQUEST FOR COMMENT: INT'L ACCOUNTING STANDARDS (2000). There are a number of important areas of disagreement between GAAP and the international standards, including differences in recognition, measurement, reporting requirements, and presentation. Financial Accounting Standards Board, *The IASC-U.S. Comparison Project: A Report on the Similarities and Differences between IASC Standards and U.S. GAAP* 41 (1999). Notably, accounting recognition criteria are currently the subject of much dispute in the Enron case, with auditors calling for revisions of the GAAP requirements. Of particular concern in the Enron case was the use of special-purpose entities that can be left off the consolidated books of a parent company as long as 3% of their capital comes from outsiders. Daniel Kadlec, *The [Enron] Spillover*, TIME, Feb. 4, 2002, at 28. The SEC may want to revisit its position on GAAP, which is frequently criticized as being too historically (rather than current market) based. Dynegy, Mirant and General Electric Co. have all boosted disclosure relating to special purpose entities in their annual reports, even in the absence of new regulation—though perhaps in anticipation of foreseeable future regulation. *See* Rachel Emma Silverman, *GE's Annual Report Bulges with Data in Bid to Address Post-Enron Concerns*, WALL ST. J., Mar. 11, 2002, at A3 (noting that "GE went out of its way to distinguish its off-balance sheet practices from Enron's"); *Accounting: Dynegy and Mirant Enhance Valuation Disclosure in Andersen-Audited 10-Ks*, ELECTRIC UTIL. WK., Mar. 18, 2002, at 8.

73. Rachel McTague, *Congress: Lieberman Calls Hearings on Enron on Senate Side to Look at Government Role*, 34 Sec. Reg. & L. 5 (BNA), at 5 (Jan. 7, 2002).

the appearance of Enron's finances, debt was transferred to undisclosed partnerships run by Enron executives in which officers had a personal stake.[74] By sharing directors with the entities it created, Enron violated the cardinal principle of corporate independence that validates special purpose entities.[75] In addition, the partnerships were used to enrich Enron executives while they were supposedly representing Enron in negotiating self-dealing transactions.[76] New rules are being considered to require independence of equity investors, a rule that Enron would have violated due to the heavy investment of its then-chief financial officer, Andrew Fastow, and his colleagues.[77]

Inflating revenue is a widespread problem, as evidenced by investigations of Computer Associates International (for questionable accounting practices that enriched top executives),[78] CMS Energy (for counting sham energy trades as revenue), Dynegy (for sham energy trades), Global Crossing (for trading fiber optic capacity in order to record sales that were never made), Halliburton (for booking cost overruns as revenue although it might not be paid for the excess), Quest Communications International (for creating trades in fiber optic capacity that had no economic value), Reliant Resources (questionable energy trades), Waste Management (falsifying earnings), and Xerox (for including future payments on existing contracts in its current revenue).[79] These companies frequently argue that they are observing the technicalities of GAAP. Nonetheless, their practices fail to disclose the true financial picture.

Moreover, recent studies show that managers are manipulating disclosure to increase their own compensation.[80] They can do this because performance-based compensation, originally conceived to better align management and investor

---

74. Enron Report, *supra* note 58, at 12, 21 (observing that on three occasions the board approved the creation of special purpose entities, which were partly owned and wholly managed by Enron executives, to do business with Enron, and that each time the presentations of these entities and their transactions were described "in light of their favorable impact on Enron's financial statements").

75. Davis, *supra* note 61, at 42. For a discussion of the role of special purpose entities in Enron's collapse, see generally Steven L. Schwarcz, *Enron and the Use and Abuse of Special Purpose Entities in Corporate Structures*, 70 U. CIN. L. REV. 1309, 1314-15 (2002).

76. Enron Report, *supra* note 58, at 26, 27 (noting conflicts arising from the substantial profits that accrued to Enron executives transacting business through the "internal Enron marketplace").

77. Davis, *supra* note 61, at 42.

78. Gaston F. Ceron, *Staying Focused: Corporate Governance May Be Everybody's Responsibility; But at Some Companies, One Person Has More Responsibility Than Others*, WALL ST. J., Feb. 24, 2003, at R7 (finding the "software maker has been bruised by criticism of its executive compensation practices . . . [and an SEC] investigation into its accounting [practices]").

79. Laura S. Egodigwe et al., *A Year of Scandals and Sorrow*, WALL ST. J., Jan. 2, 2003, at R10.

80. Charles M. Yablon & Jennifer Hill, *Timing Corporate Disclosures to Maximize Performance-Based Renumeration: A Case of Misaligned Incentives?*, 35 WAKE FOREST L. REV. 83 (2000).

interests, is now a dominant form of pay in the United States and other industrialized countries.[81] The reality is that these packages unduly favor management.[82] Performance based compensation coupled with stock option grants have increased executive compensation far out of proportion to any actual increase in economic growth.[83]

This is an important issue. Using stock option compensation to align officers' interests with those of shareholders doesn't work.[84] Any incentive to look after shareholder interests can be defeated by the possibility of stock sales, hedging through derivatives, manipulation of earnings accruals on the books, and timed disclosures of bad news. These are agency costs—incentives for self-dealing—imposed by managers on the firm and its principals, injuring firm prospects, value and credibility. In Enron, management, including directors, held off disclosure of massive losses just long enough to unload a fair quantity of their stock.[85]

Enron's auditors also failed to disclose conflicts of interest which arose from their stake in selling the special entity partnerships as investment vehicles.[86]

---

81. Stuart Weinberg, *Insiders Hedge With Zero-Cost Collars*, WALL ST. J., Aug. 7, 2002, at B5 (noting that "[m]any executive-pay packages include company shares in order to link executives' interests to the fortunes of the companies they manage").

82. Kathy M. Kristof, *CEOs Paid 70% More at Firms Under Scrutiny Accounting: Top Officers at 23 Companies Made Well More Than Average, According to a Study*, WALL ST. J., Aug. 26, 2002, at C3 (noting the " increasingly controversial practice of paying top executives with stock options that become valuable only if the company's market price rises [gives] executives the incentive to inflate profits to drive up their companies' stock prices").

83. Yablon & Hill, *supra* note 80, at 85. FASB has taken the position that disclosure of executive stock options in financial statement footnotes is sufficient, and that it is not necessary to expense stock-based compensation. Nonetheless, it recognizes that "disclosure is not an adequate substitute for recognition of assets, liabilities, equity, revenues, and expenses in financial statements." *Stock Options*, 34 Sec. Reg. & L. Rep. (BNA) (Mar. 4, 2002) (quoting IASB chair David Tweedie). Further, the International Accounting Standards Board (IASB), which has yet to set a standard for stock option compensation accounting, has noted that the real questions are whether it is an expense, and if so, why it does not appear in the income statement.

84. *See, e.g.*, Lucian A. Bebchuk et al., *Managerial Power and Rent Extraction in the Design of Executive Compensation*, 69 U. CHI. L. REV. 751, 755 (2002) (concluding that "managerial power and rent extraction are indeed likely to play a significant role in executive compensation in the United States").

85. *See* Enron Report, *supra* note 58, at 47-51 (explaining the significant of inadequate disclosure coupled with lavish stock option bonuses in that "giving Enron executives huge stock option awards, they might be creating incentives for Enron executives to improperly manipulate company earnings to increase the company stock price and cash in their options").

86. These special entity partnerships are called trust-preferred securities. Goldman Sachs pioneered this investment strategy in which a corporation formed a subsidiary (in this case they were limited partnerships) that issued preferred shares paying a fixed amount to investors; the proceeds were then loaned to the corporation and treated as an asset. Apparently a change in the accounting rules is what caused Enron to have to restate its financials to account for these

They apparently made over $1 million in fees for helping to structure the transactions, far more than from auditing the company. Enron is not the only company in which the auditing process failed to reveal impending financial disaster. Waste Management, Sunbeam, and Microstrategy are only a few of the firms in which auditors were similarly deficient.[87] Inadequate disclosure of loans or losses has surfaced as a major problem in a number of corporations now being investigated by the SEC, including Adelphia Communications (which failed to disclose $3.1 billion in loans and loan guarantees to its founder's family), Kmart (which is being investigated for its loss accounting), Tyco (which may have made undisclosed loans to its former CEO and other top executives), and Worldcom (which made more than $400 million in loans to its former CEO).[88] Ultimately these kinds of conflicts of interest affect the financial health of the corporation, and are information that should be disclosed.

II. EVOLUTIONARY GAME THEORY AND THE IMPORTANCE OF REGULATION

Prior to the spate of corporate fiascos, of which Enron was a part, deregulation was a popular proposal. Commentators frequently relied on game theory to justify deregulation of the securities industry. Properly applied, however, game theory suggests that deregulation actually will lead to the kind of corporate malfeasance that Enron illustrates. In particular, the branch of iterated game theory with repeat players known as evolutionary game theory, provides a useful lens through which to understand the problem of corporate monitoring, and to assess the role of law in structuring human interactions—in this case, the decisions made by those in charge of large publicly held corporations. Evolutionary game theory offers insights into the dynamics of two kinds of interactions that drive corporate governance: interactions between corporate insiders and investors, and interactions between directors and managers.

A. Basic Game Theory Concepts

Game theory is a way of mathematically modeling strategic interactions.[89] It attempts to simplify a social interaction in which at least two people (called the players) must choose a course of action. The result of the game is called a payoff. The players' choice of strategy will be based on their payoff, and what the players predict the others will do. Game theory, like conventional economics, attempts to predict human behavior by ignoring irrelevant details and focusing on the essence of a particular choice of behavior. Its goal is to predict which

---

partnerships. *See* John D. McKinnon, *Enron Tax Strategy Opposed by Clinton Draws Scrutiny of Government Officials*, WALL ST. J., Jan. 15, 2002, at A18.

87. *Oversight of Accountants and Auditors: SEC Proposes New Regulatory Regime, Oversight Panel for Accountants, Auditors*, 34 Sec. Reg. & L. Rep. (BNA) (Mar. 4, 2002).

88. Hubert B. Herring, *An Impossible Dream: Corporate Honesty*, N.Y. TIMES, June 16, 2002, § 3, at 12.

89. For an elegant introduction to game theory (without the mathematical notations), see BAIRD ET AL., *supra* note 6.

strategies the players are likely to choose, and a fundamental assumption is that players will prefer higher payoffs to lower, taking into account what the other players are likely to do.[90] Game theory assumes that players believe that other players will also act to optimize their payoffs.[91] A strong assumption of game theory is complete information, which means that the players know the strategies available to each player and the payoffs to every combination of strategies.[92] The combination of strategies in which no player could do better by changing strategies is known as the Nash equilibrium.[93]

The best known of the normal form games[94] is Prisoner's Dilemma, which explores whether people can be motivated to behave cooperatively. Two players must decide whether to cooperate or not; if both cooperate they each receive

---

90. *See id.* at 11 (explaining that the idea that each player will choose the best outcome in light of what the other player is likely to do—that is, a player will choose a dominant strategy over a dominated strategy—is "the most compelling precept in all of game theory").

91. *See id.* at 13 (illustrating the concept of iterated dominance through a normal form game between a pedestrian and motorist in which the available strategies are exercise care or exercise no care, where exercising care costs both players $10, an accident is certain to happen if neither exercises care, 10% certain if both exercise care, and an accident costs the pedestrian $100; if neither exercises care the motorist receives a payoff of $0 and the pedestrian a payoff of $100; if the motorist exercises no care, but the pedestrian exercises care, the motorist's payoff is $0 and the pedestrian's is -$110; because not exercising care is a dominant strategy for the motorist, the pedestrian will predict that is the strategy the motorist will adopt and therefore will not exercise care either). This model is highly stylized, in that it assumes no legal rule to shift liability, the motorists' indifference to causing harm, and an absence of harm to the motorist; nonetheless, it predicts that in a regime of no liability, the motorist will have too little incentive to take optimal care to minimize accidents. *Id.* at 14. A change in the legal rules—the game structure—will change the incentives of both players. *Id.* at 14-15.

92. *Id.* at 312. Stag hunt and prisoner's dilemma are examples of games with complete, though imperfect, information.

93. John Nash, *Equilibrium Points in N-Person Games,* 36 PROC. NAT'L ACAD. SCI. 48-49 (1950). In a single play prisoner's dilemma game, for example, there are four possible strategies: cooperate-cooperate, cooperate-defect, defect-defect, and defect-cooperate. *See, e.g.*, ALEXANDER J. FIELD, ALTRUISTICALLY INCLINED? THE BEHAVIORAL SCIENCES, EVOLUTIONARY THEORY, AND THE ORIGINS OF RECIPROCITY 2 (2001) (discussing the concept of Nash equilibria). Defect-defect is the Nash equilibrium in single or finitely repeated prisoner's dilemma games. *Id.* at 2. It is not, however, a pareto optimal solution, in that the good of both players taken as a whole would be cooperate-cooperate, which would yield six total points rather than the maximum of five for defect-cooperate. However, from the individual player's standpoint, defection is strictly dominant, in that it results in a higher payoff regardless of the other player's strategy. *See id.* at 2 (discussing the concept of strict dominance).

94. Normal form games are those consisting of the following three elements: players, strategies and playoffs. BAIRD ET AL., *supra* note 6, at 311. Two-by-two, or bimatrix games, consist of two players, each of whom has a small number of strategies, represented by a box of four squares, the payoffs of which are listed by convention with the row player's first, and the column player's second. *Id.* at 303.

three points; if both defect (by not cooperating), they receive one point each; but if one player defects and the other cooperates, the defector receives five points and the cooperator receives zero (the sucker's payoff).[95] More total points will be garnered if they both cooperate, but from either player's standpoint it is better to defect.[96] Under these circumstances, the logical thing to do is defect because no matter what the other player's choice (cooperate or defect), the payoff will be better for the defector.[97] The Nash equilibrium, the point where no one has an incentive to deviate from the chosen strategy, is defection by both players.[98] Always Defect is the strategy of rational self-maximizers. This is true whether the game is played just once or for a set number of repetitions.

There are two inter-related games that this Article is concerned about: the issuer/investor game and the corporate governance game. The interaction between the shareholders and the issuer can be modeled as a prisoner's dilemma game where the overall payoffs for the group would be enhanced by cooperation, but where the temptation to defect results in payoffs that are far from optional because the issuer's defect position of grabbing the investors' money will tempt investors to play their defect position of keeping their money under a mattress. Thus, the investors' defect position will predominate.[99] It is a prisoner's dilemma (a kind of collective action problem) because, in a well-functioning, fraud-free market, the insiders and the investors would both be better off if the investors invested and the issuer pursued profits rather than creating the illusion of profits.

In the corporate governance game, the players are the corporation's managers and directors. The role that directors are assigned as firm monitors has dynamics similar to the well-known agency-regulated firm game, where for the firm defection means law evasion and for the regulator defection means punitive enforcement.[100] In the corporate governance game, defection by management

---

95. Martin A. Nowak et al., *Cooperation Versus Competition*, 56 FIN. ANALYSTS J. 13, 16 (2000).

96. This example is taken from BAIRD ET AL., *supra* note 6, at 33.

97. A similar game, Wolf's dilemma, in which twenty people sit in cubicles, with their fingers on buttons. After ten minutes each person will get $1000 as long as no one pushes a button. If someone pushes a button before time, that person gets $100 and everyone else gets nothing. Because there is a small chance that someone will push the button, logic impels each person to try to be the first pusher. *See* DOUGLAS R. HOFSTADTER, METAMAGICAL THEMAS: QUESTING FOR THE ESSENCE OF MIND AND PATTERN (1985).

98. *See* ANATOL RAPOPORT & ALBERT M. CHAMMAH, PRISONER'S DILEMMA (1965).

99. *See, e.g.*, BAIRD ET AL., *supra* note 6, at 46 (describing lending as a two-by-two normal form game in which, absent law, no lending activity will take place). As Baird and his co-authors conclude, unless there are strong legal protections assuring investors that their money is not going to be stolen, no money will be invested. *Id.*

100. IAN AYRES & JOHN BRAITHWAITE, RESPONSIVE REGULATION: TRANSCENDING THE DEREGULATION DEBATE 54, 55 (1992) ("Business regulation is often modeled as a game between two players—the regulatory agency and the firm."). This inescapably oversimplifies the matter, but it has advantages of clarifying the dynamics of the interaction. *Id.* at 55. John Scholz models business regulation as a prisoner's dilemma, where the motivation of the firm is to decrease

means evading full disclosure when presenting the board with information upon which it must take action as a means of decreasing costs of compliance.[101] Defection by directors means refusal to ratify management decisions, or in the extreme case, firing management, motivated by attempts to achieve maximum compliance with shareholder protections.[102]

In a single prisoner's dilemma game, one would expect both sides to defect, that is, for management to attempt to hide information from the board that would contradict management's desired outcome, and the board to insist on full disclosure regarding the impact of management proposals on shareholders. Presumably, that insight is the reason for having a board in the first place. As the

---

regulatory costs and the motivation of the regulator is to maximize compliance. John T. Scholz, *Cooperation, Deterrence, and the Ecology of Regulatory Enforcement*, 18 LAW & SOC'Y REV. 179, 192 (1984).

101. To the detriment of shareholders, not all managers are exclusively motivated by their own self-interest. As Ayres & Braithwaite noted:

> [s]ome corporate actors will comply with the law if it is economically rational for them to do so; most corporate actors will comply with the law most of the time simply because it is the law; all corporate actors are bundles of contradictory commitments to values about economic rationality, law abidingness, and business responsibility. Business executives have profit-maximizing selves and law-abiding selves, at different moments, in different contexts, the different selves prevail.

AYRES & BRAITHWAITE, *supra* note 100, at 19.

102. Another way of visualizing the game between the directors and managers is a variation on the prisoner's dilemma, the hawk-dove game. Like prisoner's dilemma, the hawk-dove game is a two-by-two normal form game. *See* BAIRD ET AL., *supra* note 6, at 45. For a description of this game and its relationship to the generation of oppressive norms, see Amy Wax, *Expressive Law and Oppressive Norms: A Comment on Richard McAdams's "A Focal Point Theory of Expressive Law,"* 86 VA. L. REV. 1731, 1732 (2000). This game also emphasizes interactions where the players can benefit from cooperation but have incompatible claims to resources. *Id.* The choices are to be assertive against the other player (play hawk) or to be submissive (play dove). Each player's first choice is to play hawk while the other plays dove. The second choice is for both the players to play dove. Third choice is to play dove while the other player plays hawk. Both players want to avoid playing hawk, because conflict is inevitable, and the costs of losing exceed the benefits of winning. Neither player knows what the strategy of the other player will be. Although the role assigned to directors is nominally that of hawk (they are supposed to monitor, after all), in practical terms they serve at the managers' will. On the other hand, the directors do have the power (though rarely exercised) to fire management in a disagreement over policy. Therefore, there is some fluidity in the roles of the players. This suggests that in repeated games, a player that was hawk may shift to dove and vice versa. Unlike prisoner's dilemma, the hawk-dove game has no single Nash equilibrium in which the strategy for each player is the best reply to itself. *Id.* at 1733. But like prisoner's dilemma, infinite repetition has the potential to change the outcome to a cooperative strategy. *See, e.g.*, Robert J. Axelrod & Robert O. Keohane, *Achieving Cooperation Under Anarchy: Strategies and Institutions*, 38 WORLD POL. 226, 231 (1985) ("What is important for our purposes is not to focus exclusively on Prisoner's Dilemma per se, but to emphasize the fundamental problem that it (along with [hawk-dove and stag hunt]) illustrates.").

following section demonstrates, however, repeat interactions change the outcome to a cooperative strategy. This has advantages for the public good in the issuer/investor game, and disadvantages for the public good in the corporate governance game.

B. Evolutionary Game Theory

Evolutionary game theory combines the idea of evolutionary interactions with a branch of mathematical economics to demonstrate that human beings are social animals for whom reciprocity is as important as competition.[103] Competition and cooperation together drive the engines of evolution and economy.[104] Evolutionary game theory demonstrates the conditions under which cooperation will emerge despite a constant urge to defect.[105] Rather than strict competition, an alternative—and more accurate—vision is that of coevolution, with its emphasis on the importance of initial conditions for cooperation to flourish and for the gains from trade to be shared.[106]

The twist that evolutionary game theory adds to the Prisoner's Dilemma game is that always defecting is not the most successful strategy in games repeated indefinitely.[107] An infinite number of strategies is possible for the

---

103. Evolutionary game theory demonstrates that in paired interactions, rather than strictly acting to maximize profit, people assess their partner's projected outlook and act in response. These results are explained by evolutionary biologists as conferring an evolutionary advantage to a species that lives in groups and can expect future encounters both within the group and with other groups. Evolutionary biologists theorize that the way human beings think is a product of evolutionary history, and depends not only on brain structure but also on adaptive responses to the environment (including other human beings). *See generally* Nowak et al., *supra* note 95.

104. *See* STUART KAUFFMAN, AT HOME IN THE UNIVERSE 240 (2000) (noting that "[a]n economy, like an ecosystem, is a web of coevolving agents" and explaining the analogy between evolution and economy in terms of complexity theory).

105. A standard Darwinian economic argument is that whatever economic institutions survive are presumptively efficient. *See* Mark Roe, *Chaos and Evolution in Law and Economics*, 109 HARV. L. REV. 641, 641 (1996). This is mistaken because, as Roe explains, chaos, path dependencies, and the concept of local equilibria assure that nothing about survival implies "superiority to untried alternatives." *Id.* at 643.

106. One of the key insights of complexity theory is the importance of initial conditions for self-organization, and the role of positive feedback mechanisms. *See* RICHARD SOLE & BRIAN GOODWIN, SIGNS OF LIFE 299 (2000) (arguing that complex systems, such as traffic patterns, internet use, and by extension, trade transactions, need a "distributed and adaptive set of rules, always in direct interaction with the internal web dynamics [in order to] . . . effect cooperation (and not conflict) among users"). *Cf.* Guido Calabresi, *The Pointlessness of Pareto: Carrying Coase Further*, 100 YALE L. J. 1211, 1229 (1991) (castigating the use of Pareto "to hide the inevitability of distribution issues").

107. *See* Gerald S. Wilkinson, *Reciprocal Food Sharing in the Vampire Bat*, 308 NATURE 181, 181-84 (1984) (observing that a past blood donator will receive blood from a prior recipient, but a past refuser will not, and that bats seem to be good at keeping score).

iterated game.[108] These repeated games of Prisoner's Dilemma illustrate a major anomaly that undermines the paradigm of self-interested profit maximizing competition in classic utility theory. Self-interest may include cooperation and retaliation, neither predicted by utility theory, because such a cooperative strategy may be more effective in evolutionary terms.[109] Maynard Smith theorized that a strategy is evolutionarily stable if no differing strategy can invade a population of repeat players.[110] Evolutionary game theory thus provides a bridge between biology and economics by explaining the interaction of cooperative and competitive behavior.[111]

Testing this theory, Robert Axelrod devised a series of computer tournaments to confront populations of strategies in repeated games of Prisoner's Dilemma.[112] In these tournaments, a strategy called "tit-for-tat" seems to rule.[113] Tit-for-tat begins by cooperating and then responds in kind to whatever the other player did the last time.[114] The advantage of tit-for-tat is "its combination of being nice, retaliatory, forgiving and clear."[115] In a computer tournament where nasty strategies (always defect), nice strategies (always cooperate) and retaliatory strategies (tit-for-tat) played repeatedly against each other, tit-for-tat prevailed.[116]

---

108. *See* Karl Sigmund, *Automata for Repeated Games*, *in* EVOLUTION AND PROGRESS IN DEMOCRACIES 335, 336 (Johann Gotschl ed. 2001) (noting in that in the context of iterated games there are too many Nash equilibria for the concept to provide a solution).

109. *See* J. Maynard Smith & G. R. Price, *The Logic of Animal Conflict*, 246 NATURE 15, 15 (1973) (arguing that in the context of animal conflicts, strategies are hard-wired into genetic modes of behavior, so that randomly meeting players will play the game according to their genetic programming; the more successful strategies in a population will predominate because randomly meeting individual strategies will be tested against each other, and if one strategy is consistently more successful, it will dominate the population). John Maynard Smith, an evolutionary biologist, puzzled over this result, and postulated that it was linked to the courting behavior of animals, which rarely fight to the death in mating contests. Maynard Smith modeled these contests as two-by-two normal form games known as hawk-dove games. *Id.* Maynard Smith drew on Darwinian selection and postulated that this result must be a Nash equilibrium that is stable over time. He termed the biological equivalent of the Nash equilibrium an "evolutionary stable strategy" (ESS) in which natural selection causes animals to behave instinctively with similar strategies. *See id.* (exploring why animals do not generally fight to the death and concluding that in repeat games a cooperative strategy may prevail). Notably, although an ESS is a Nash equilibrium, not all Nash equilibria are ESS. *See* FIELD, *supra* note 93, at 140 (explaining that the best counter to an ESS must be itself). An ESS cannot be successfully invaded by any other strategy, and it may consist of "more than one strategy in stable proportions." *Id.*

110. J. MAYNARD SMITH, EVOLUTION AND THE THEORY OF GAMES (1982).

111. *See* MATT RIDLEY, THE ORIGINS OF VIRTUE 54 (1996) (explaining developments in the evolutionary theory of behavior).

112. ROBERT AXELROD, THE EVOLUTION OF COOPERATION (1984).

113. WILLIAM POUNDSTONE, PRISONER'S DILEMMA (1992).

114. AXELROD, *supra* note 112, at 6.

115. *See id.*

116. *See* RIDLEY, *supra* note 111, at 61 (detailing the Axelrod computer tournament). This

Although the success of a tit-for-tat strategy in repeat encounters might appear to be a good explanation for human cooperation, it is not, because while cooperation is encouraged by frequent repetition, so is retaliation. If two tit-for-tatters are playing repetitively, they will cooperate until one defects; but once defection has begun, a downward spiral of retaliation begins. That is the downside of tit-for-tat and may explain the incessant tribal feuds in places like Israel, Ireland, the Balkans, and Afghanistan.[117] It also suggests the problems that a regulatory scheme based primarily on punishment may encounter.[118]

This result sparked a great deal of research into reciprocal behavior in the animal world.[119] Although tit-for-tat is practiced by some animals (notably bats and reef fish) it is strikingly absent from most of the animal kingdom, an empirical result that led to further refinements of the Axelrod tournaments.[120] Tit for Tat only prevails where the conditions for the contest are stable. Where conditions were made more random,[121] in an attempt to simulate real world conditions, a new strategy, a random Tit for Two Tats, prevailed for a while but then permitted Always Cooperate to prevail—a situation ripe for exploitation by Always Defect. Thus, this result does not explain human cooperation either,

---

is possible because Prisoner's Dilemma is not a zero sum game. Even a small minority of tit-for-tatters can hold its own against a majority of Always Defectors and eventually prevail because, for the most part, the individual tit-for-tatters will receive slightly less of a payoff than an Always Defector. However, in the few games against other tit-for-tatters this is more than made up for, and the more the population of tit-for-tatters grows, the larger the joint payoff. *See* Sigmund, *supra* note 108, at 338 (describing the Axelrod tournaments).

117. It may also explain the importance of reputation and other signaling devices. Picking the right partner to bargain with is crucial. In a world of defectors, tit-for-tat cannot take hold unless it can find other cooperators. Playing Prisoner's Dilemma with strangers heightens the importance of trustworthiness signals. One of the reasons people advertise trustworthiness, through facial expressions, actions, and reputation, is to identify people who are not opportunists and attract them. ROBERT H. FRANK, PASSIONS WITHIN REASON (1988) (explaining the role of emotions in advertising and identifying trustworthiness).

118. *See* AYRES & BRAITHWAITE, *supra* note 100, at 25 (observing that a "mostly punitive policy . . . fosters an organized business subculture of resistance to regulation. . . .").

119. *See* RIDLEY, *supra* note 111, at 61.

120. *See, e.g.*, Sigmund, *supra* note 108, at 339 (introducing the notion of random error).

121. Robert Sugden proposed an asychronous game (for a strategy called "Contrite Tit for Tat"—which I call Tit for Two Tats), in which a state "good" or "bad" is assigned to each player; the player chooses to cooperate or defect, and is assigned a new state depending on both the player's prior move and prior state. ROBERT SUGDEN, THE ECONOMICS OF RIGHTS, COOPERATION AND WELFARE (1986). Sigmund explains the strategy:

> If my opponent was in state g or both players were in b, I achieve state g if I have played C, or else b. But if I was in state g while the opponent was in b, then I enter state g, no matter whether I have played C or D. Hence, if I play C, I always achieve state g. If I play D, however, I will get into state g only if, in the previous round, I was in g and opponent in b.

Sigmund, *supra* note 108, at 340-41.

because it is an unstable downward spiral.[122] Nor does it provide much guidance for how to structure a regulatory game to achieve an optimal balance of cooperation and competition.

A more stable strategy is Pavlov, a strategy that repays its partners in kind (like for Tit for Tat), forgiving occasionally (like Tit for Two Tats), but with a nasty streak that lets it exploit Always Cooperate.[123] A Pavlov player starts with cooperate and defects in the next round if the opponent's move was different.[124] If both players defected in the prior round, Pavlov tries cooperating; if the opponent cooperated when Pavlov defected, Pavlov defects again. If Pavlov cooperated while the opponent defected, Pavlov's next move is defection. This makes for a stable strategy, because if one player erroneously chooses to defect, cooperation is reestablished in two rounds.[125] Pavlov's flaw, however, is that in an Always Defect environment, it cannot prevail.[126]

A better strategy is Firm-but-Fair, which is slightly nicer than Pavlov, in that it "cooperates with cooperators, returns to cooperating after a mutual defection, and punishes a sucker by further defection, but unlike Pavlov it continues to cooperate after being the sucker in the previous round."[127] The success of Firm-but-Fair strategy is evolutionarily stable for Prisoner's Dilemma and is a possible explanation for the evolution of human cooperative behavior.[128] This is because, when it is surrounded by a population of Always Defect players, Firm-but-Fair acts like Tit for Tat in that it cannot be duped for more than one round and does not cooperate again until the loss is made up, but it is better because it does not need a cooperative population to prevail.[129] But it does more. It illustrates how one might view a system that balances minimum regulatory interference with optimal payoffs.

The studies on the evolution of cooperation may have salutary implications for the investor/issuer game, because it implies that cooperation may become a stable strategy over time. On the other hand, it is not so wonderful for the corporate governance game, because it suggests that directors will tend to be captured by the managers they are supposed to monitor.[130] The tendency toward

---

122. RIDLEY, *supra* note 111, at 77.

123. M.A. Nowak et al., *The Arithmetics of Mutual Help*, 272 SCI. AM. 50-55 (1995).

124. *See* Sigmund, *supra* note 108, at 340 (describing Pavlov strategy as representing "the simplest rule of learning: to repeat something if and only if it was satisfying the last time").

125. *Id.*

126. *Id.*

127. RIDLEY, *supra* note 111, at 80.

128. *See* Jonathan Bendor & Pitor Swistak, *The Evolutionary Stability of Cooperation*, 91 AM. POL. SCI. REV. 290, 290 (mathematically demonstrating Axelrod's intuition that efficient strategies have an evolutionary advantage).

129. Sigmund, *supra* note 108, at 341.

130. *See* AYRES & BRAITHWAITE, *supra* note 100, at 63 (modeling regulatory capture and demonstrating that capture changes the payoff matrix so that "firm defection will not lead to the joint defection equilibrium, but to the firm defect:agency [sic] cooperate equilibrium"). This shift is equally likely in the management/director interactions.

cooperation is further illustrated by a game with human players, the Ultimatum Game (and its variations).

C. A Game with Human Players: The Ultimatum Game

Reciprocity only works in small groups where individuals can keep track of past generosity or defection. While these are the kinds of interactions that directors and managers have with each other (in the corporate governance game), their interactions with their investors (in the issuer/investor game) are impersonal. Yet, people often cooperate with people from whom they cannot expect reciprocal favors, and obey rules that are essentially unenforceable, such as pooper-scooper rules for dog owners, and smoking bans in buildings.[131] Although it would be rational to be a free-rider in a complex society, most people do not free-ride. The Ultimatum Game and the Group Exchange Game illustrate this tendency toward cooperation.

Both traits predicted by evolutionary game theory—cooperation and retaliation—can be demonstrated in the Ultimatum Game. Neither of these traits is predicted by neoclassical law and economics. Cooperation in the Ultimatum Game is illustrated through a human tendency toward fairness, in the sense of sharing gains from trade. This tendency toward fairness appears irrational, unless one thinks in terms of the selfish gene theory, and the evolutionary importance of cooperation.[132] The tendency to retaliate is similarly apparently "irrational," causing people to incur costs to themselves in order to punish a defector.[133]

In the Ultimatum Game, two players are isolated from each other.[134] Player A is given a sum of money and told to propose a share with player B. Player B knows the amount at stake, and if player B refuses the offer, neither player gets anything. If Player B accepts, the money is shared accordingly. No bargaining

---

131. RIDLEY, *supra* note 111, at 180.

132. *See* W. D. Hamilton, *The Genetical Evolution of Social Behavior*, 7 J. THEORETICAL BIOLOGY 1, 1-52 (1964) (interpreting the selfless behavior of ants in caring for their sisters' offspring as the behavior motivated by the selfish genes of the ants, because chances of genetic survival are greater by acting selflessly than by procreating individually). Conversely, even relationships that have always been assumed to be altruistic, such as the mother's nurture of her child in utero, have an element of competition. *See* D. Haig, *Genetic Conflicts in Human Pregnancy*, 68 Q. REV. BIOLOGY 495, 495-31 (1993) (describing the hormonal conflicts between mother and child over blood sugar levels as an example of diverging genetic interests).

133. This paradox of how harmony prevails over selfishness in biological terms is resolved by understanding that for each self-interested gene that would be only too happy to individually self-maximize (in the form of cancer, for example, the paradigmatic mutiny of selfish cells), there are many others that will combine to suppress it. *See* EGBERT LEIGH, ADAPTATION AND DIVERSITY (1971) ("It is as if we had to do with a parliament of genes: each acts in its own self-interest, but if its acts hurt others, they will combine together to suppress it.").

134. For a description of the Ultimatum Game, see Karl Sigmund et al., *The Economics of Fair Play*, 286 SCIENTIFIC AM. 83 (2002).

is permitted. The logical amount for Player A to offer is any amount greater than zero; if strict rationality were driving the decision, Player B would accept. But in fact, the most common offer made by Player A is around one-half of the original sum, demonstrating a strong tendency toward sharing gains. If the offer is much less than one-half, Player B almost invariably refuses.[135] This demonstrates a willingness not only to punish the selfish offeror, but also to incur costs in doing so (because neither player receives anything on refusal).[136] The results of this game are remarkably consistent across cultures.[137]

Cooperation, retaliation, and a tendency toward fairness are context dependent; however, altering the structure of the Ultimatum Game changes these results. For example, if the right to be Player A is earned (say, by high scoring in a test of knowledge or winning the right in a contest), Players A tend to be less generous, and Players B more accepting of low offers.[138] One of the dangers of an extremely hierarchical "rank and yank" system such as the one at Enron,[139] is that it may promote self-dealing behavior. Players A also tend to be less generous if Players B must accept the offer (the Dictator Game).[140] Anonymity also has an effect. This has implications for the issuer/investor game, because investors in large publicly held corporations are not known personally to corporate insiders. If Player A's identity is protected even from the experimenter, 70% offer nothing.[141] If several Players B compete to accept the offer, Players A may offer a smaller percent.[142] This also has negative implications for generosity of issuers toward their investors. If the tendency to fairness is undermined by anonymity and hierarchy while self-dealing behavior

---

135. *See id.* (noting that, on average, two-thirds of offers are between 40 to 50%; only one-quarter are for less than 20% of the pot; more than one-half of responders reject offers of less than 20% of the pot).

136. Acts of punishment appear to violate the conventional economist's notion of rationality because rational actors are supposed to ignore sunk costs. *See* Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L. REV. 1551, 1562-63 (1998) (referring to sunk costs as letting bygones be bygones). An evolutionary explanation of the desire for punishment is that it disciplines self-maximizers who refuse to cooperate in much the same way as the selfish gene theory proposes that the body's cells cooperate to attack mutinous cancer cells. *See* RIDLEY, *supra* note 111, at 180.

137. RIDLEY, *supra* note 111, at 141-42. *See* Dapluna Lewinson-Zamir, *The Choice Between Property Rules and Liability Rules Revisited: Critical Observations from Behavioral Studies*, 80 TEX. L. REV. 219, 227 & n.33 (2001) (listing studies showing the "wide divergence of experimental results from theoretic predictions . . . in different countries and cultures").

138. *Id.* at 140; Sigmund et al., *supra* note 134, at 84.

139. *See* Malcolm Gladwell, *The Talent Myth*, NEW YORKER, July 22, 2002.

140. Surprisingly, in the dictator game, in which Player A proposes a division of the pot, but Player B has to accept, results show that although Player A offers tend to be less generous division than in the Ultimatum Game, most Players A offer a significant percentage of the pot, even if players identities are hidden. KROBOTKIN & ULEN at 1136.

141. RIDLEY, *supra* note 111, at 140-41.

142. Sigmund et al., *supra* note 134, at 84.

is promoted, efficient norms may not prevail. Thus, the structure of the game as well as the tendency to fairness are important.

D. Structuring Games: The Role of Law

Game theory helps to explain why deregulation is misconceived and how collusive behavior between corporate managers and directors can emerge even in the absence of explicit agreement. As Ellikson discovered, social norms may evolve in the absence of law among small communities of repeat players.[143] But once the community becomes larger and more anonymous, as in the issuer/investor game, or rules and sanctions become less clear, evolutionary game theory emphasizes the importance of initial conditions and argues strongly for a role of law to promote investor confidence.[144]

1. Initial Conditions: The Importance of Structure for the Emergence of Markets.—Game theory sees the function of law as structuring the interactions between players who may have diverging interests and asymmetrical information.[145] It focuses on what the players in a particular game with specified objectives observe, what they can infer, and what they are likely to do.[146] The structure of the game determines the players' payoffs, their strategies, and the possibility and number of Nash equilibria. Single prisoner's dilemmas are quite different from repeated transactions with the same players. Confronted with defect as a first move, tit-for-tat and Pavlov keep retaliating. Even firm-but-fair spends a lot of time retaliating. None of this is likely to be socially optimal.

Evolutionary game theory suggests that one role of law is to alter the players' incentive structures in order to make socially valuable transactions more likely. For example, in a loan contract in the absence of legal liability for default, a transaction is unlikely to take place even though it would be better for society if it did, because the lender's optimal strategy would be not to lend.[147] Law thus

---

143. ROBERT C. ELLICKSON, ORDER WITHOUT LAW 52-64 (1991).

144. *See* Nowak et al., *supra* note 95, at 21 (noting that "cooperation is greater in a sedentary than in a mobile population" because "[d]efectors can thrive in an anonymous crowd").

145. Game theory is a way of studying the strategic behavior of interdependent individuals with divergent economic interests. Eric Talley, *Interdisciplinary Gap Filling: Game Theory and the Law*, 22 LAW & SOC. INQUIRY 1055, 1057 (1997) (noting that although "law plays an integral role in shaping and regulating the interaction between players who possess possibly divergent interests and beliefs," its effect depends on "the players' individual and common understanding (or lack thereof) of the existence, content, and applicability of legal rules") (reviewing BAIRD ET AL., *supra* note 6). Asymmetric information games are those in which the players are not completely informed about each others' payoffs (but do know that they are incompletely informed). "Incomplete information is the central problem in game theory and the law." BAIRD ET AL., *supra* note 6, at 33.

146. For an elegant description of the fundamental concepts of game theory, see Stephen W. Salant & Theodore S. Sims, *Game Theory and the Law: Ready for Prime Time?*, 94 MICH. L. REV. 1839, 1845 (1996) (book review).

147. Assigning numbers to this idea, if we assume a gain from the loan of $10 to be shared

helps to transform a game with suboptimal equilibria into a game with optimal solutions. Similarly, in impersonal market transactions, investors who believe that the game is rigged and sanctions unavailing may withhold their money from the market. An often-cited justification for the mandatory disclosure regime and the antifraud provisions is that of encouraging investors to participate in the market.

2. The Role of Sanctions.—One purpose of legal rules is to sanction and thereby raise or lower the costs for certain behavior.[148] Voluntary acceptance of rules that promote participants' objectives is undoubtedly preferable to sanctions as an economic solution to achieving cooperative behavior.[149] It is certainly cheaper. But game theory explains that stabilizing cooperative interactions requires would-be defectors to face the threat of sanctions and "that those who are charged with identifying defectors and carrying out such sanctions be sufficiently motivated to do so."[150] In repeated interactions, informal norms of reciprocity may emerge, but only if participants know each other and expect that defection will be met with retaliation at the next iteration of the game.[151] Enforcement by third parties—courts, for example—may be necessary in the absence of a small community of repeat players.[152]

3. Solving Information Asymmetries.—Solving information asymmetry problems is important in game theory as it is in classical economics. The assumption of common knowledge in game theory is a strong one. If one of the players is mistaken about the other player's past move, this can lead to an endless cycle of defection in a population of tit-for-tatters.[153] Legal rules can help clarify which ambiguous moves count as defection.[154] Bargaining that takes place over

---

equally, without legal liability for default, if the lender makes no loan, its payoff is 0, and the borrower's payoff is 0; if the lender loans $100 and the borrower defaults, the lender is out $110, and the borrower gains $110; if the borrower repays the loan, the lender's maximum payoff is $5, and the borrower's is $5. This example is from BAIRD ET AL., *supra* note 6, at 46. In the presence of liability for default, however, the lender's maximum payoff is $ 5 whether the borrower defaults or not, which is preferable to the $0 the lender would receive if no transaction took place. *Id.*

148. Law and economics goals are to "promote the most efficient allocation of resources—maximizing wealth and minimizing costs." Mark R. Fondacaro, *Toward an Ecological Jurisprudence Rooted in Concepts of Justice and Empirical Research*, 69 UMKC L. REV. 179, 181 (2000).

149. Edward F. McClennon, *Pragmatic Rationality and Rule*, *in* EVOLUTION AND PROGRESS IN DEMOCRACIES 181, 183 (Johann Gotschl ed. 2001) (arguing that a commitment to rules is instrumentally rational as a way of solving coordination problems).

150. *Id.* at 209-10 n.55.

151. *Id.* at 200.

152. *See id.* (observing that increased reporting and punishment of defectors yield increased cooperation if others in the community, who are not necessarily co-players, also retaliate).

153. BAIRD ET AL., *supra* note 6, at 174.

154. *See* Avery Katz, *The Strategic Structure of Offer and Acceptance: Game Theory and the Law of Contract Formation*, 89 MICH. L. REV. 215, 236 (1990) (noting that "modern research on bargaining has revealed that additional theoretical refinements, concerning information and the

time, rather than instantaneously, or that takes place in the absence of information (about the players' situation, the possible outcomes of a strategy, about whether the other player is a cooperator or a defector), often have more than one Nash equilibrium.[155] Because legal rules affect the way information is transferred between parties and the timing of its transfer and can direct the parties toward socially desirable solutions, disclosure and liability rules are important for market players.

a. The problem of nonverifiable information.—Legal rules and third party enforcement mechanisms cannot solve all the problems of information asymmetry. Information may be inaccessible to third parties. Where information is imperfect and mistakes possible (through misreading signals, for example), tit-for-tat may devolve into constant defection. For example, a legal rule prohibiting misrepresentation gives a seller, whose Nash equilibrium might otherwise be to lie about the number of apples in a box, the incentive to disclose the correct number. The number of apples is verifiable, both by the buyer and by the court; however, not all information that the parties need is verifiable. The quality of apples, for example, or their suitability for gift baskets, may not be so easily determined. Similarly, investment contracts (whether explicit or implicit) need some mechanism of ensuring that both parties act optimally even though the court may know less than they do. The possibility of third party mistakes makes adjudication a strategic choice in this situation.[156] This may be one explanation for the consternation caused by plaintiffs' strike suits.

b. Reputation and signaling as a solution to information asymmetries.—Even proponents of deregulation acknowledge the informational asymmetries between managers (who have access to firm information) and investors (who do not), on which mandatory disclosure rules are justified. Investor demand for information coupled with the use of firm quality signaling devices such as outside auditors, reputable investment bankers, managerial stock purchases and dividend payment, are supposed to suffice. Because much of the information for sound investment decisions is unverifiable, however, reputation and signaling were intended to separate the stars from the dregs.

Signaling is a way that people with nonverifiable information can convey that information through the way they act.[157] For a while, at least, people believed that informational quality signaling devices included the use of outside auditors for corporate financial statements, hiring as underwriters reputable investment bankers, managerial stock purchases, and stock dividend payments, for example.

In a wider community without interconnecting social and business ties—the investing public, for example—the problem of nonverifiable information looms

---

timing of the parties' strategies, need to be imposed on bargaining models if useful results are to be obtained").

155. *See id.* at 235, 237 (discussing the effects of timing and information on Nash equilibria).

156. Richard H. McAdams, *A Focal Point Theory of Expressive Law,* 86 VA. L. REV. 1649, 1699 (2000).

157. *See* BAIRD ET AL., *supra* note 6, at 123-24 (explaining that the use of warranties may be a way of signaling high quality goods, a signal that legal rules requiring warranties may impede).

large. Eric Posner asserts that a wide variety of conduct serves to signal potential partners about the cooperative intentions of a player.[158] Posner argues that merchants investing in expensive office space, correct grooming, social speech, gift-giving, and similar conduct all signal what he calls the discount rate—the player's cooperativeness—to future transactional partners.[159] If, indeed, such conduct is meant to convey one's cooperative nature, however, it is easy to see how such a wide variety of behaviors could be misinterpreted (or even strategically adopted to fake other players out). Similarly, the adoption of one of the formerly Big Five accounting firms did little to ensure the quality of Enron's or WorldCom's financial statements. Signaling from Enron's outside auditors provided no information regarding the quality of the firm because its auditors were themselves conflicted, as the recipients of large amounts of consulting revenue stemming from advice on setting up the very partnerships that were hiding the company's losses).

Reputation also failed to prevent defection: Enron was one of the most widely followed and well-respected companies in the United States. There is some reason to be skeptical about the efficacy of analyst reports as signaling devices. Even in large companies that are followed by analysts, buy recommendations far outweigh sell recommendations, perhaps as a result of analyst overoptimism). Enron, a large, widely followed company elicited no cautionary statements in the financial press despite impending financial disaster. Because Enron executives were selling off their holdings and because Enron created thousands of special purpose entities, analysts now uniformly agree that there were numerous danger signals that should have been reported in the financial press. The Enron debacle and large number of corporate defections provide reasons to doubt the efficacy of reputation and signaling devices in impersonal markets.

Reputational concern may check opportunistic behavior, but its importance will depend on the frequency of similar transactions, the length of the time horizon, and how profitable the transaction.[160] Enron and the other large publicly held corporations that have defected in the issuer/investor game had good reputations. Enron, Global Crossing, and WorldCom were widely followed companies. Moreover, when the situation becomes an end game, the returns to reputation may not be sufficient to prevent opportunistic behavior.[161] The threat of bankruptcy may thus diminish the importance of reputation in the decisional calculus.

4. Channeling Behavior in Coordination Games.—Law may have yet another function: coordinating players' moves. Coordination problems are persistent in social interactions even where people share an interest in an

---

158. *See* ERIC POSNER, LAW AND SOCIAL NORMS 19-24 (2000).

159. *Id.* at 23.

160. *See* MILGROM & ROBERTS, *supra* note 20, at 139 (discussing the ways in which concern for reputation may prevent defection, but also noting that its value depends on how often it will prove useful).

161. *See id.* at 266 (discussing the end game problem).

efficient outcome, and an equally important function of law is that it may serve as a focal point for individuals to coordinate their actions.[162] The famous stag hunt game, for instance, attributed to Jean Jacques Rousseau, requires the players to coordinate their efforts to achieve optimal payoffs.[163] The stag hunt game involves two players and two strategies: hunt hare or hunt stag. It takes two to hunt stag, and half a stag is better than a whole hare. Both players would prefer to hunt stag together, but would prefer to hunt hare alone over the possibility of hunting stag alone, and neither knows what the other will do. The Nash equilibria are for both to hunt stag, both to hunt hare; or a random mix of the two.[164] This is known as a complete but imperfect information coordination game. The players know the payoffs and the available strategies, but neither knows what the other will do. It is in both the players' best interests to hunt stag, but without communication, they are unlikely to do so.

The issuer/investor game may be seen as a form of stag hunt. Both issuers and investors will be better off if they pool their resources and "hunt stag" together. Yet, due to fear of the other's defection, each may instead hunt hare alone. Law may provide the necessary signal for issuers and investors to pool their resources. Line item disclosure rules and periodic reporting requirements coordinate the types of disclosure that will be required and provide some way of comparing companies through their disclosure. As Richard McAdams explains, the "law provides a focal point around which individuals can coordinate their behavior."[165]

Although a Nash equilibrium is a self-enforcing strategy that cannot be improved upon as long as the other players are following their prescribed strategies,[166] some games, such as the stag hunt game and the hawk-dove game,

---

162. *See* McAdams, *supra* note 156, at 1652 (arguing that even a sanctionless proclamation may cause people to change their expectations of what others will do, thus influencing behavior).

163. BAIRD ET AL., *supra* note 6, at 36.

164. *Id.*

165. McAdams, *supra* note 156, at 1651. Although McAdams characterizes his game theory analysis as a form of rational choice theory, he imports a distinctly behavioral economic view of rationality. *See, e.g., id.* at 1662 & 1663 n.38 (claiming an empirical basis for his claims and acknowledging that his account is an expanded form of rational choice in that "individuals are rationally exploiting features of their environment, even if the perceptions of those features are not themselves determined solely by rationality."). The concept of focal point in coordination games is illustrated by a game where a group of people is given the project of meeting in New York City on a given day without being permitted to communicate the time or place of meeting. *See* THOMAS C. SCHELLING, THE STRATEGY OF CONFLICTS (1963). In an empirical test, most people given this task met at noon under the clock in Grand Central Station. *Id.* Although there were an infinite number of Nash equilibria (every time and location), the Grand Central clock at noon acted as a focal point. *See* BAIRD ET AL., *supra* note 6, at 39.

166. *See* Talley, *supra* note 145, at 1059 (explaining that the concept of Nash equilibrium is important because "it delineates the behavior in which rational, self-interested actors would plausibly engage").

may have multiple Nash equilibria.[167] If there are two or more equilibrium points, law serves a channeling function, enabling the parties to choose between them (by enabling them to predict which of these strategies will prevail).[168] A common example is choosing on which side of the street to drive. Both right and left are equilibrium points, as is random right and left, but without some coordinating force, there is no a priori way to determine whether the players will choose random sides, right or left.[169]

E. Undermining the Role of Law in the Issuer/Investor Game

So much for games. How does this apply to real life? Enron provides a useful lens, although the past two years have provided a bumper crop of other illustrative examples.[170] These corporate defections have followed a period of questionable enforcement of legal rules. A misguidedly "hands-off" approach to regulation, coupled with obstacles to private enforcement, crippled third party enforcement mechanisms and produced a climate conducive to defection. Deregulation was the order of the day, built on free-market arguments that "[i]f disclosure is worthwhile to investors, the firm can profit by providing it."[171]

1. Globalization and Deregulation.—Deregulation has been gathering momentum over the past two decades.[172] Two key—and interrelated—developments have spurred deregulation of the capital markets. The first is the internet, with its potential for instantaneous communication with investors around the world. The Internet increasingly is being used as a way to reach

---

167. *See id.* 1059 & n.6 (describing a simple coordination game of two players with three Nash equilibria in which neither player has any incentive to change strategy: choosing whether to drive on the right or left side of the road, where one solution is that both players choose the right side of the road, the second solution is that both players choose the left; and the third solution is to randomly choose to drive on the right or left).

168. *See* McAdams, *supra* note 156, at 1659 (explaining that law serves "to coordinate predictions . . . to identify the one course of action that their expectations of each other can converge on" by helping the players to "'mutually recognize' some unique signal that coordinates their expectations of each other").

169. *Id.*

170. WorldCom is being investigated for spectacularly overstating its assets. Kurt Eichenwald & Simon Romero, *Turmoil at WorldCom: The Decision Making*, N.Y.TIMES, June 27, 2002, at A1 (noting overstatements of WorldCom assets by $3.8 billion). Tyco, GE, IBM, Morgan Chase, all use similar "off-balance-sheet" partnerships to manage their finances. PNC Financial Services Group is being investigated by the SEC for using off-balance sheet entities to hide underperforming assets, as is Dynegy. *Bush Doctrine*, N.Y.TIMES, June 16, 2002, at 12. Managers at Global Crossing and Lucent, like those at Enron, similarly cashed out while withholding information that their companies were going down in flames. ImClone's former chief executive officer (CEO) has been indicted for insider trading. *Id.*

171. Easterbrook & Fischel, *supra* note 5, at 682.

172. *See* Steinberg, *supra* note 31, at 348 (observing "widespread accommodations for both domestic and foreign issuers").

investors, inform them about investment opportunities, as well as to effect trades. Web pages for issuers and securities brokerage firms are proliferating. Frequently, hyperlinks exist between them. There is no reason to suppose that their audience will be limited by national boundaries.

At the same time, globalization, with its accompanying increase in international competition, is forcing a new transnational perspective on many financial institutions, including the securities industry. Foreign issuers have dramatically increased their presence in the U.S. markets, through registered offerings, private placements, and American Depository Receipts.[173] Fifteen percent of the listed corporations on the New York Stock Exchange are now foreign issuers, and American investors now plough over $5 trillion into foreign securities.[174] U.S. issuers have similarly sought out foreign markets.

Reluctant to forgo the diversification benefits to American investors, the SEC has dramatically liberalized its approach to disclosure requirements for foreign issuers.[175] Furthermore, sensitive to the charge that it is creating a dual regime of easy access to U.S. markets for foreigners, coupled with much more onerous disclosure duties for domestic issuers, the SEC has also revisited many of its disclosure requirements for domestic issuers in an attempt to streamline its requirements and lessen the burden on regulated entities.[176] As a result, both foreign and domestic issuers have a much greater choice about the level of

---

173. American Depository Receipts are ownership interests in a particular number of securities held by a U.S. bank or trust company. *See generally* Joseph Velli, *American Depository Receipts: An Overview*, 17 FORDHAM INT'L L.J. 38 (1994).

174. U.S. Statistical Abstract 838 (1999).

175. Foreign issuers offering securities in the United States must register using forms F-1, F-2 and F-3 that parallel the registration forms S-1, S-2, and S-3 required of United States issuers. The SEC has lightened their disclosure burdens, however, by permitting financial statements to be prepared according to the accounting principals in their locale, as long as they explain material variations from U.S. GAAP requirements. *See* JAMES D. COX ET AL., SECURITIES REGULATION: CASES AND MATERIALS 327 (3d ed. 2001) (discussing the streamlining of disclosure requirements for foreign issuers in the United States). Foreign issuers are also excused from disclosing line-of-business information, and have reduced disclosure obligations with respect to management compensation, and management transactions with the issuer. *Id.* Regulation S facilitates offshore offerings.

176. For example, the accommodations for domestic issuers include:

> [T]he adoption of the "accredited purchaser" principle under Regulation D, thereby dismantling the mandatory disclosure framework in offerings made solely to accredited purchasers and relegating such investors to private redress under federal law exclusively to the Section 10(b) anti-fraud remedy; shortening the holding period to resell restricted securities to one year and permitting unrestricted resales by nonaffiliates of the subject issuer after a two-year holding period; [and] authorizing extensive incorporation by reference in registered offerings by less than premier issuers pursuant to the shelf [sic] registration rule and SEC Form S-3 . . . .

Steinberg, *supra* note 31, at 348-49.

disclosure they will have to provide.[177]

In addition to the deregulatory trend of the SEC, Enron illustrates other issues posed by the absence of government regulation. Its on-line energy trading systems functioned without any government oversight because over-the-counter derivatives are not regulated, by either the SEC or the Commodity Futures Trading Commission (CFTC).[178] In testimony before the Senate Governmental Affairs Committee, Professor Frank Portnoy said that Enron employees "used dummy accounts and rigged valuation methodologies to create false profit and loss entries for the derivatives Enron traded."[179] A bill currently before the Senate would give the CFTC the authority to regulate energy derivatives, rolling back the commodity law exemption for over-the-counter derivatives.[180] However, the CTFC chair, James Newsome, told the Senate Energy and Natural Resources Committee that he remains committed to a deregulated futures industry.[181] As a result, it remains unlikely that much will be done to regulate these vehicles.

2. Litigation Barriers.—In theory, defecting corporate managers face punishment, not only from government enforcement efforts, but also from private litigants. Private enforcement potentially is available through two major legal avenues: the federal securities laws and state law fiduciary duty obligations. The securities laws place disclosure obligations on corporate management in the form of line item disclosure obligations, reporting obligations, and liability under the antifraud provisions. State fiduciary duties arise from agency principles and include the duty of care[182] and the duty of loyalty.[183] From these basic precepts, the duty of candor and the duty to disclose arise.[184]

The SEC has traditionally refrained from addressing state law corporate governance standards.[185] Two important concerns are the primacy of federal

---

177. *See* Alan R. Palmiter, *Toward Disclosure Choice in Securities Offerings*, 1999 COLUM. BUS. L. REV. 1, 3 (1999) (mandatory nature of securities regulation has been eroding for the last twenty years, due to rethinking the "twin tenets that have animated securities regulation for more than half a century—namely, manager informational shirking and investor helplessness").

178. *Enron Case Highlights Policy Question, How Much to Regulate OTC Derivatives*, 34 Sec. Reg. & L. Rep. (BNA), at S-7 (Mar. 4, 2002) [hereinafter *OTC Derivatives]*.

179. *Id.* (quoting Frank Portnoy).

180. S. 1951, 107th Cong. § 1(c) (2002).

181. *OTC Derivatives*, *supra* note 178, at S-8.

182. RESTATEMENT (SECOND) OF AGENCY § 379 (1958) (calling for agents to exercise the care and skill "standard in the locality for the kind of work . . . [and] any special skill" the agent has).

183. *Id.* § 387 (requiring that the agent work "solely for the benefit of the principal in all matters connected with his agency").

184. *See* Lawrence A. Hamermesh, *Calling Off the Lynch Mob: The Corporate Director's Fiduciary Disclosure Duty*, 49 VAND. L. REV. 1087, 1115 (1996) (describing the genesis of the Delaware duty of candor); RESTATEMENT (SECOND) OF AGENCY § 381 (1958) (requiring that an agent has the duty to disclose all matters relating to the agency).

185. *See, e.g.*, *In re* Franchard Corp., 42 S.E.C. 163 (1964) (declining to mandate disclosure of the directors' dereliction of duty for anything less than "[o]utright fraud or reckless indifference"

litigation in the securities area and the fear of inconsistent rulings regarding fiduciary duty—corporate governance—issues. Federal securities law requires initial, transaction-specific and periodic disclosure, and its antifraud provisions require that any disclosures made be truthful and not misleading.[186] In contrast, most state corporate governance law requires disclosure only in the event of shareholder action.[187] Under both federal and state law, claims are based on allegations that the corporation made a false or misleading statement. However, in both areas, litigation is becoming more difficult.

a. Statutory barriers.—The Private Securities Litigation Reform Act of 1995 (PSLRA)[188] raised procedural hurdles aimed at reducing perceived abuses of federal class action securities suits,[189] while the Securities Litigation Uniform Standards Act of 1998 (SLUSA)[190] preempted the ability of plaintiffs to take their securities claims to state court (where the hurdles might be lower).[191] Together,

---

amounting to "total abdication" of their role). Under PSLRA, two types of state disclosure actions are preserved: for misrepresentations involving a buy-back or going-private transaction and for misrepresentations in connection with a tender offer, merger, or exchange offer. This is the so-called "Delaware Carve-Out." *See The Securities Uniform Standards Act of 1997-S 1260: Hearings on S. 1260 Before the Comm. on Banking, Housing and Urban Affairs, Subcomm. on Securities*, 105th Cong. 69, 73 (1998) (noting that "the Delaware courts can resolve these claims . . . in a matter of days or weeks") (statement of John F. Olson).

186. *See, e.g.*, Jennifer O'Hare, *Director Communications and the Uneasy Relationship Between the Fiduciary Duty of Disclosure and the Anti-Fraud Provisions of the Federal Securities Laws*, 70 U. CIN. L. REV. 475, 479-82 (2002) (discussing the legal framework of the federal disclosure and anti-fraud provisions).

187. Delaware is an exception, requiring that even if shareholder action is not requested, any disclosure that is made must be truthful and not misleading. *See* Malone v. Brincat, 722 A.2d 5, 9 (Del. 1998).

188. Pub. L. No. 104-67, 109 Stat. 737 (1995).

189. The PSLRA additionally provides, among other things, for imposition of penalties on the plaintiff if the lawsuit was found to be frivolous, proportional rather than joint and several liability, a safe-harbor for firms' forward-looking statements, and choosing lead counsel. *Id.*

190. 15 U.S.C. § 77k(b) (2000). *See, e.g.*, Lander v. Hartford Life & Annuity Ins. Co., 251 F.3d 101 (2d Cir. 2001) (holding that SLUSA applied to a class action alleging misrepresentations in the sale of annuity contracts, which activated removal); *In re* Bank Am. Corp. Secs. Litig., 263 F.3d 795 (8th Cir. 2001) (explaining that Congress was funneling class-action securities litigation into the federal courts).

191. The purpose of SLUSA was to discourage plaintiffs from circumventing the PSLRA by filing in state rather than federal court, to encourage disclosure of forward-looking information, and to establish uniform rules for securities class actions. H.R. CONF. REP. NO. 105-803, at 13 (1998). SLUSA preempts state law entirely where it applies, so that not only do plaintiffs lose their right to bring securities claims in state courts, they also lose their right to litigate state claims in federal court through supplemental jurisdiction. O'Hare, *supra* note 186, at 489. SLUSA preempts state law securities class actions based on "misrepresentations or omissions of material fact in connection with the purchase or sale of covered securities; or . . . that the defendant used or employed any manipulative or deceptive device in connection with the purchase or sale of covered

these statutes attempted to protect businesses from plaintiffs' attorneys and coercive settlements.[192] Proponents of PSLRA argued that much pre-PSLRA plaintiffs' securities litigation was frivolous, but that because of high legal defense fees, even non-meritorious suits were settled.[193]

The courts, however, were already well-equipped to handle this problem. Federal courts are no strangers to motions to dismiss. Difficult discovery issues and the handling of sensitive materials are a frequent phenomenon in federal courts.[194] Judges can and should manage discovery.[195] There is little empirical evidence supporting these statutes: the 1995 Act was part of the general tort reform movement, in large part instigated by the insurance, hi-tech venture, and accounting lobbies,[196] and the 1998 Act was based on the dubious and controversial claim that litigants were circumventing the intent of Congress by

---

securities." SLUSA, Pub. L. No. 105-353, §101(a)(1) & (b)(1), 112 Stat. 3227. SLUSA preempts "covered class actions," which it defines as suits "on a representative basis" or on behalf of "more than 50 persons." *Id.* This is a significant barrier to litigation, because most securities fraud actions are prohibitively expensive for individual investors, so that they tend to be brought as class actions. The "Delaware Carve out," however, exempts shareholder derivative actions and claimed breaches of fiduciary duty, even when they involve securities disclosures. SLUSA §16(d) (exempting certain covered class actions). *See* Arlia v. Blankenship, 234 F. Supp. 2d 606, 612-13 (S.D. W. Va. 2002) (derivative claim for breach of fiduciary duty exempted from SLUSA under § 16(d)); Dediger v. Tallman, 2000 WL 268309 (Del. Ch. Feb. 25, 2000) (referring to § 16(d) as the "Delaware carve-out"). A shareholders' derivative action is brought on behalf of the corporation for harm to the corporation. *See* Roberta Romano, *Corporate Governance in the Aftermath of the Insurance Crisis,* 39 EMORY L.J. 1155, 1156-57 (1990) (discussing the difference between direct and derivative claims).

192. *See* 141 CONG. REC. S17,933-04, S17,954 (daily ed. Dec. 5, 1995) (statement of Mr. Dodd). The purpose of the PSLRA was to curb perceived abuses in private securities actions. *See* H. R. Conf. Rep. No. 104-369, at 32 (1995). PSLRA also provides a statutory safe harbor for forward looking statements. Securities Act of 1933 § 27A(c), 15 U.S.C. § 77z-2(c) (2000); Securities Exchange Act of 1934 § 21E(c)(1)(A)(I), 15 U.S.C. § 78u-5(c) (2000). *See* H.R. REP. NO. 104-369, at 43 (1995) ("Understanding a company's own assessment of its future potential would be among the most valuable information shareholders and potential investors could have about a firm.") (remarks of former SEC Chair Richard Breeden).

193. *See, e.g.*, Ahiq Ali & Sanjay Kallapur, *Securities Price Consequences of the Private Securities Litigation Reform Act of 1995 and Related Events*, 76 ACCT. REV. 431-60 (2001) (noting arguments that PSLRA was needed to deter nuisance suits).

194. *See* Joel A. Seligman, *The Merits Do Matter*, 108 HARV. L. REV. 438, 438 (1994) (arguing that pre-PSLRA courts were well equipped to dismiss nonmeritorious suits).

195. *See, e.g.*, Crawford-El v. Britton, 523 U.S. 574, 598 (1998) (observing that the trial judge has "broad discretion to tailor discovery narrowly and to dictate the sequence of discovery").

196. *See* Hillary A. Sale, *Heightened Pleading and Discovery Stays: An Analysis of the Effect of the PSLRA's Internal-Information Standard on '33 Act and '34 Act Claims*, 76 WASH. U. L.Q. 537, 556 (1998) (explaining the genesis of the PSLRA as part of the general tort reform movement, in Congress's "Contract with America").

taking their claims to state court.[197]

PSLRA requires the plaintiff to "state with particularity facts giving rise to a strong inference that the defendant acted with the required state of mind."[198] Upon the filing of a motion to dismiss, PSLRA provides an automatic stay of discovery.[199] As Hillary Sale noted, this combination is "outcome determinative and, if strictly applied, virtually impossible to meet" in private securities litigation.[200]

Although previous pleading standards (specifically Rule 9(b) of the Federal Rules of Civil Procedure) also required pleading fraud with particularity, pre-PSLRA plaintiffs were able to obtain needed internal company information through the Federal Rules of Civil Procedure's liberal discovery provisions.[201] In enacting PSLRA, Congress expected to deter strike suits[202] by requiring dismissal if plaintiffs fail initially to plead not only the reason or reasons why an alleged misstatement was misleading when made but also facts giving rise to a strong inference that the defendants acted with scienter.[203] Without discovery, however, plaintiffs are unable to obtain documents such as board meeting notes, internal audit documents, or internal memoranda—the very documents most plaintiffs need to document their claims.[204] As a result, not only strike suits but also meritorious fraud claims may be quelled.[205]

---

197. Mark Klock, *Lighthouse or Hidden Reef? Navigating the Fiduciary Duty of Delaware Corporations' Directors in the Wake of* Malone, 6 STAN. J. L. BUS. & FIN. 1, 31 (2000).

198. 15 U.S.C. § 78u-4(b)(2) (2000).

199. *See* Securities Act of 1933 § 27A(f), 15 U.S.C. § 77z-2(f) (2000); Securities Exchange Act of 1934 § 21D(6)(3)(B)), 15 U.S.C. § 78u-4b(3)(B) (2000).

200. Sale, *supra* note 196, at 538.

201. *Id.* (noting the change wrought by PSLRA's pleading and stay rules). Professor Sale explains that "in order to meet the common-law pleading standards developed by the Ninth and Second Circuits prior to the [PSLRA], plaintiffs needed access to internal company information" which they obtained "by engaging in discovery and then repleading their complaints." *Id.* at 539. Former-President Clinton expressed a similar concern in his veto message, stating that the bill would have "the effect of closing the courthouse door on investors who have legitimate claims." 141 CONG. REC. H15,214, H15,214 (daily ed. Dec. 20, 1995) (statement of President Clinton). Then-SEC Chairman Arthur Levitt similarly warned that the legislation would undercut investors' rights. *Id.* at H15,220.

202. A strike suit is defined by one business dictionary as a "derivative action, usu[ally] based on no valid claim, brought either for nuisance value or to obtain a settlement." A HANDBOOK OF BUSINESS LAW TERMS 579 (Bryan A. Garner ed. 1999).

203. 1934 Act § 21D(b)(1)-(2).

204. *See, e.g.*, *In re* Silicon Graphics, Inc. Secs. Litig., 970 F. Supp. 746, 767-68 (N.D. Cal. 1997) (finding allegations of insider trading together with dates and contents of negative internal reports insufficient absent titles, dates, authors, recipients, contents and sources of reports).

205. *See* Sale, *supra* note 196, at 564 (contending that "the Reform Act is likely to allow only the more flagrant and obvious cases of securities fraud to proceed past a motion to dismiss, while being overinclusive in its elimination of cases where it is more difficult to identify, and therefore to plead, fraud" which "is likely to result in unredressed fraud").

The worry impelling passage of PSLRA was that a typical securities fraud case was filed not because the plaintiff had discovered fraud, but instead because there had been a sudden drop in stock price.[206] However, only a small fraction of companies whose stocks plummet experience such filings.[207] Moreover, the pre-PSLRA courts were well aware of this factor and used the particularity requirements to screen cases that were merely responses to a decrease in price.[208] Nor is there any a priori reason to believe that securities fraud allegations are beyond the courts' competence to parse.[209]

Moreover, solicitude for the inability of businesses to defend themselves from such suits appears unfounded. The costs of insurance and litigation defense were argued as grounds for the bill, but these costs were never substantiated.[210]

---

206. *See* John W. Avery, *Securities Litigation Reform: The Long and Winding Road to the Private Securities Litigation Reform Act of 1995*, 51 BUS. LAW. 335, 339 (1996) (discussing the arguments of reform advocates, who contended that any drop of 10% of the stock price prompted fraud claims).

207. *See, e.g., Securities Litigation, 1994: Hearings on H.R. 417 Before the Subcomm. on Telecomm. and Fin. of the House Comm. on Energy and Commerce*, 103d Cong. 118, 119 (1994) (testimony of Donald C. Langevoort, Lee S. and Charles A. Speir Professor of Law, Vanderbilt Univ. School of Law) (arguing for reform, but suggesting that there was no meaningful correlation between stock-price drops and fraud claim filings); 103d Cong. 267 (1994) (testimony of Leonard B. Simon, attorney); 141 CONG. REC. S17,933-04, S17,951 (daily ed. Dec. 5, 1995) (statement of Mr. Bryan) (citing University of California study demonstrating that of 589 stocks that dropped 20% in price within a five-day period, only 3% were sued). Even the bill's proponent, Senator Domenici, could say no more than that 21% of securities fraud cases were filed within forty-eight hours of a drop in price. 141 CONG. REC. S17,965-03, S17,968 (daily ed. Dec. 5, 1995) (statement of Mr. Domenici) (citing study by National Association of Securities and Commercial Law Attorneys that found that 21% of fraud cases were filed within forty-eight hours of a price drop).

208. *See* Sale, *supra* note 196, at 544 (discussing courts' response to the fear that nonmeritorious suits were being filed simply because the stock price dropped, in particular the requirement that plaintiffs plead facts to show that the difference was attributable to fraud).

209. This is not a suggestion that courts are unconcerned about the difficulties presented. Courts certainly have expressed concern about adjudicating competing interests in the absence of statutory or regulatory guidance. *See, e.g., In re* Time Warner Inc. Sec. Litig., 9 F.3d 259, 263 (2d Cir. 1993) (noting the "inevitable tension between two powerful interests" and complaining that "the adjudication process is not well suited to the formulation of a universal resolution of the tensions" in the absence of statutory or regulatory guidance); DiLeo v. Ernst & Young, 901 F.2d 624, 627 (7th Cir. 1990) (recognizing that "only a fraction of financial deteriorations reflect fraud"). Notably, however, courts frequently express dismay at resolving complex issues in a wide variety of settings. *See, e.g.*, Erica Beecher-Monas, *Blinded by Science: How Judges Avoid the Science in Scientific Evidence*, 71 TEMP. L. REV. 55 (1998) (discussing the courts' rhetoric of dismay at having to decide the admissibility of expert testimony and opining that courts are quite capable of making such determinations).

210. For example, in the "Joint Explanatory Statement" the managers of the House and Senate, in support of their recommendation, proffered only the statement of "the general counsel of an investment bank" regarding high discovery costs. *See* Conference Report on H.R. 1058, with Joint

The argument that businesses were being held up for ransom and forced to settle meritless litigation was wholly unsupported by empirical data. Subsequently, empirical studies have shown that most pre-PSLRA cases settled and that settlements were tied to the merits.[211] In addition, pre-PSLRA cases routinely settled for well under 20% of the potential investor losses.[212]

Although the number of securities fraud cases has not fallen post-PSLRA,[213] this may be due to an increase in meritorious securities fraud cases that can obtain the necessary factual basis without discovery. It says nothing about the effect of PSLRA on meritorious cases that cannot obtain such information without discovery.[214] A number of changes have occurred in the kinds of cases filed post-PSLRA. As the empirical study of Mukesh Bajaj and his co-workers demonstrated, the number of accounting fraud (including revenue restatements), improper accounting practices and improper revenue recognition cases filed increased, while the percentage of material omissions cases decreased somewhat.[215] Perhaps these cases are less difficult to substantiate pre-discovery. One interesting and unanticipated result is that the settlement rate decreased post-PSLRA from 57.6% within four years of filing to 26% within four years of filing.[216] Thus, it would appear that survival of the PSLRA dismissal process has made litigants more willing to bring their cases to trial. However, it is difficult to draw firm conclusions from these data because of the considerable variation in the data over time.[217] Another unanticipated consequence of impeding meritorious lawsuits is that it may reduce incentives for honest disclosure.[218]

Statement of Conference Comm., H.R. REP. No. 104-369, 104th Cong., 1st Sess., *reprinted in* 27 Sec. Reg. & L. Rep. (BNA), at 1893 (Dec. 1, 1995).

211. *See* Stephen P. Marino & Renee D. Marino, *An Empirical Study of Recent Securities Class Action Settlements Involving Accountants, Attorneys, or Underwriters*, 22 SEC. REG. L.J. 1159 (1994) (demonstrating that settlements are tied to the merits of the case).

212. *See* Mukesh Bajaj et al., *Securities Class Action Settlements: An Empirical Analysis*, working paper, *available at* http://securities.stanford.edu/research/studies/20001116_SSRN_Bajaj.pdf, at 12 (Nov. 16, 2000) (noting that the mean settlement ration to potential investor loss amounts varied from 7.1% for cases filed in the Fourth Circuit to 21.9% for cases filed in the Tenth Circuit).

213. *See id.* at 3 (noting that the number of federal cases filed had "reached an all-time high of 248 filings" in 1998).

214. *See* Perino, *supra* note 4, at *22 (acknowledging that it is impossible to know whether meritorious suits are being chilled and observing that "while it is difficult to assess the claim that there is more fraud now than there was prior to the PSLRA, the other explanations for the apparent increase in filings appear to be inadequate"); Ribstein, *supra* note 2, at 17 (arguing that "reduced liability risk may have encouraged fraudulent or shirking behavior in marginal situations where defrauding insiders or lax auditors had persuaded themselves that the likelihood of detection was low . . . [which] argues for reversing some aspects of the PSLRA").

215. Bajaj et al., *supra* note 212, at 4.

216. *Id.* at 5.

217. *Id.* at 13.

218. *See* Seligman, *supra* note 194 (arguing that corporate officers and advisors will have

This is substantiated by some evidence that shareholders consider PSLRA harmful.[219]

b. Judicial barriers.—PSLRA and SLUSA are not the only new barriers to securities actions. In addition to statutory changes in the legal landscape, the Supreme Court, in *Central Bank of Denver v. First Interstate Bank of Denver*,[220] held that private fraud actions under the Securities Exchange Act section 10(b) and SEC Rule 10b-5 cannot be brought under an aiding and abetting theory.[221] Instead, they must be based on primary liability—the statements must be attributable to the defendant.[222] The statements in reports filed with the SEC are signed by the directors; hence, their liability for material misstatements or omissions is not changed by the *Central Bank* decision. Nor is primary liability for secondary actors such as accountants and lawyers changed. They are still responsible for statements that are attributable to them.[223]

What the Court did change, however, was the importance of being able to pursue claims against the primary actors—the very thing that PSLRA subsequently limited. Prior to *Central Bank*, courts had widely accepted the viability of secondary liability.[224] The courts have extended the rationale of *Central Bank* to conspiracy liability as well as aiding and abetting.[225] Although *Central Bank* makes it clear that the statements attributable to a firm outsider, like an accountant or a lawyer, may still be the source of primary liability, it is

---

fewer incentives to disclose if it becomes more difficult to bring meritorious actions).

219. *See* Ali & Kallapur, *supra* note 193 (studying stock price changes as a result in announcements relating to PSLRA and concluding that the evidence demonstrates shareholder concern that PSLRA may have harmful effects). Ali and Kallapur examined the results of prior studies that had concluded that stockholders considered PSLRA beneficial, and found that "the timing of multiple confounding events makes the interpretation of these daily returns ambiguous . . . [and] additional analyses . . . are largely inconsistent with their interpretation." *Id.*

220. 511 U.S. 164 (1994).

221. *Id.* at 191.

222. *Id.*

223. The Court did not delineate the kinds of activities that would result in primary liability for secondary defendants. *See id.* at 177. Instead, it merely observed that in some circumstances, accountants, lawyers and banks, could be primarily liable for material misstatements on which investors rely:

> Any person or entity, including a lawyer, accountant, or bank, who employs a manipulative device or makes a material misstatement (or omission) on which a purchaser or seller of securities relies may be liable as a primary violator . . . assuming *all* of the requirements for primary liability under Rule 10b-5 are met.

Id.

224. *See, e.g.*, Jill E. Fisch, *The Scope of Private Securities Litigation: In Search of Liability Standards for Secondary Defendants*, 99 COLUM. L. REV. 1293, 1297 (1999) (observing that *Central Bank* "came with little warning—courts and commentators had widely accepted the validity of aiding and abetting liability").

225. *See, e.g.*, Dinsmore v. Squadron, 135 F.3d 837, 838 (2d Cir. 1998); *In re* GlenFed, Inc. Sec. Litig., 60 F.3d 591, 592 (9th Cir. 1995).

less clear where primary liability ends and secondary liability begins.[226] This means, at least in the Second and Eleventh Circuits, that a secondary actor cannot be held liable for clients' disclosure statements, even if they participated in crafting those statements.[227]

These impediments have the unfortunate result of chilling deterrence of many of the abuses that Enron illustrates.[228] As it stands now, there is a void in the protection of defrauded shareholders and the integrity of the mandated disclosure system, especially where those primarily liable are insolvent.[229] Outside professionals are expected to serve gatekeeper roles in the disclosure system.[230] The question asked after the savings and loan debacle: "Where were the lawyers and accountants?" is equally apt in the Enron and WorldCom. Without the threat of secondary liability, the answer is far more likely to be "asleep at the wheel." Sarbanes-Oxley did little to remedy the problem. The SEC adopted a reporting rule for attorneys in response to Sarbanes-Oxley, under which lawyers must report wrongdoing up the corporate ladder and if no corporate response is forthcoming, withdraw.[231]

Securities law is not the only source of mandatory disclosure regulation; state law corporate governance statutes also require disclosure of some information that a firm's managers might otherwise prefer to keep to themselves.[232] Most

---

226. *See* Fisch, *supra* note 224, at 1300 (noting that liability for one's own representations is primary rather than secondary liability).

227. *See, e.g.*, Ziemba v. Cascade Int'l, Inc., 256 F.3d 1194, 1205 (11th Cir. 2001) (neither law firm nor accounting firm could be liable for statements unless they are "publicly attributable" to them at the time of the investment decision); Wright v. Ernst & Young LLP, 152 F.3d 169, 175 (2d Cir. 1998), *cert. denied*, 525 U.S. 1104 (1999) (outside auditor can incur liability only for those statements attributable to it). In the Third Circuit, on the other hand, a lawyer who significantly participates in drafting documents may become primarily liable as an author of those documents. *See* Klein v. Boyd, [1997-1998 Transfer Binder] Fed. Sec. L. Rep. (CCH) P 90,136, 90,137 (3d Cir. Feb. 12, 1998), *vacated on grant of reh'g*, No. 97-1143, 1998 U.S. App. LEXIS 4121, at *1 (Mar. 9, 1998) (securities lawyer's participation in drafting client disclosure documents may result in primary liability even without attribution).

228. *Cf.* Steinberg, *supra* note 31, at 350 (noting that the "federal courts, most particularly the Supreme Court, have also been influential during the past twenty-five years in restricting investor access to redress"). Although Professor Steinberg suggests that this may have had the "concomitant effect of encouraging capital formation," he provides no evidence of such encouragement, and he is unequivocal that "[i]nvestor protection has been diminished." *Id.* at 350-51.

229. Seligman, *supra* note 194, at 456.

230. *See* Fisch, *supra* note 224, at 1314 (explaining that "the securities disclosure system is premised upon the supposition that outside professionals will be involved in the disclosure process . . . as a substitute for greater supervision by government regulators").

231. *See* News Release, SEC Adopts Attorney Conduct Rule Under Sarbanes-Oxley Act, 2003 WL164827 (SEC Jan. 23, 2003).

232. *See* Merritt B. Fox, *Required Disclosure and Corporate Governance*, 62 LAW & CONTEMP. PROBS. 113, 114 (1999) (noting the dual regime). Full disclosure is a pre-requisite for shareholder ratification of director actions. *See, e.g.*, Smith v. Van Gorkom, 488 A.2d 858 (Del.

states require full disclosure in connection with shareholder action. There are very few occasions for such action, however: shareholders vote for directors, major corporate changes, and occasionally to ratify otherwise questionable director actions. Delaware is one of the few states that has explicitly imposed a duty of candor on directors requiring truthful disclosure to shareholders even in the absence of requested shareholder action.[233] In *Malone v. Brincat*,[234] shareholders alleged that the directors of Mercury Finance Company, a Delaware corporation, knowingly disseminated false financial information in required periodic disclosures over a period of four years, aided and abetted by the firm's auditors, at the end of which period the firm lost virtually all its value.[235] Reluctant to usurp the federal authority over the securities markets, the Chancery Court dismissed, finding that there was no request for shareholder action.[236] Although it affirmed the dismissal, the Delaware Supreme Court permitted the plaintiffs to refile, holding that deliberately misinforming shareholders violates directors' fiduciary duties, even in the absence of a request for shareholder action.[237] Thus, under Delaware state law, a false statement knowingly made will subject a director to liability, whether or not the statement was made in connection with a request for shareholder action.[238] This may deter conscious wrongdoing (and subsequent lying to cover it up).

The real problem, however, is oversight. Although the firm is to be managed "by or under the direction of" the board of directors,[239] the courts have been reluctant to impose monitoring duties on the board.[240] If directors make

---

1985) (requiring full disclosure for shareholder ratification of director actions).

233. *See, e.g.*, Potter v. Pohlad, 560 N.W.2d 389, 389 (Minn. Ct. App. 1997) (declining to impose duty of candor in absence of request for shareholder action).

234. 722 A.2d 5 (Del. 1998).

235. *Id.* at 8.

236. *Id.* (holding that if the release of inaccurate information into the marketplace was unconnected with a Delaware corporate governance issue, the claim was only viable under federal law).

237. *Id.* at 10.

238. *Id. But see* Hamermesh, *supra* note 184, at 1173-74 (contending that directors should not be liable under state fiduciary duty law for statements that do not elicit shareholder approval).

239. DEL. CODE ANN. tit. 8 § 141(a) (1991) (providing for that deviate from these norms); N.Y. Bus. Corp. L. § 701 (McKinney 1986) (accord); REVISED MODEL BUS. CORP. ACT §8.01 (3d ed. 1994) (accord).

240. *See* JAMES D. COX ET AL., CORPORATIONS 195 (1997) (noting that "requirements that the directors be attentive and reasonably informed are procedural in nature; the substantive requirement is that their decision have a 'rational basis'. . . [b]ut public policy considerations . . . have caused the courts not to apply these standards rigorously"). The few cases that do impose oversight duties do so in the context of banking, financial and insurance firms, which arguably have a higher fiduciary obligation. *See, e.g.*, Francis v. United Jersey Bank, 432 A.2d 814 (N.J. 1981). *But see* Brane v. Roth, 590 N.E.2d 587 (Ind. Ct. App. 1992) (noting that directors of grain cooperative breached their duty of care in failing to adequately hedge in grain market or supervise management); Robinson v. Watts Detective Agency, Inc., 685 F.2d 729 (1st Cir. 1982) (holding

egregiously bad decisions because they were asleep at the wheel,[241] to what extent are they entitled to judicial deference? In the Enron situation, for example, there appear to have been egregious monitoring failures by the board.[242] To keep judges from second-guessing director decisions that turn out, in hindsight, to be bad decisions, the business judgment rule protects directors from liability for foolish decisions.[243] Although *Smith v. Van Gorkom*[244] imposed a process requirement that directors obtain expert advice and inform themselves about management proposals, it did little to ensure that directors do anything other than go through the motions of informing themselves.[245] Even if the Enron directors had demanded the deal sheets that were supposed to inform them of intra-firm conflicts, there is no assurance that they would have reached a different disclosure decision or done anything to prevent the conflicts (since even when apprised of conflicted transactions they waived their own rules). Moreover, although *Caremark*[246] (at least, in dicta) appears to enforce a duty to ensure that information flows to the directors, it does not mandate that the directors either establish a compliance program (although they must examine and discuss the issue) or that they act upon the information that they receive.[247] The Enron directors had a sophisticated compliance program; they simply either ignored its requirements or ignored the information that they had obtained.[248]

Moreover, a director's duty to disclose has been significantly curtailed by a

---

that officer-directors of security agency breached their fiduciary duty by neglecting the business).

241. *Cf. Boards of Directors: Primary Responsibility for Recent Corporate Scandals Rests with the Board of Directors According to Panel Members at ABA Meeting*, 17 CORP. COUNS. WKLY. 258 (Aug. 21, 2002) (reporting the remarks of Neil Minow that the boards were "at the center of the perfect storm").

242. *See* Ribstein, *supra* note 2, at 3-7 (discussing oversight failures and noting that Enron's Special Committee acknowledged the board's monitoring failures).

243. *See* MODEL BUS. CORP. ACT § 8.31 (directors have no liability except for actions taken in bad faith, without reasonable belief, without adequate information, or unless they failed to exercise oversight for an extended period); BAINBRIDGE, *supra* note 14, at 296.

244. 488 A.2d 858, 890 (Del. 1985) (holding directors had failed to disclose the "fact that the Board had no reasonably adequate information indicative of the intrinsic value of the Company").

245. *But see* BAINBRIDGE, *supra* note 14, at 281 (arguing that *Van Gorkom* "created a set of incentives consistent with the teaching of literature on group decisionmaking" by encouraging "inquiry, deliberation, care, and process"). The major consequence of *Van Gorkom*, however, appears to have been a "full employment act" for investment bankers and other experts rather than a genuine search for critique.

246. *In re* Caremark Int'l Inc. Derivative Litig., 698 A.2d 959 (Del. Ch. 1996) (outlining an obligation of directors to set up a law compliance program to meet their monitoring duties).

247. Rather, the Chancery Court held that because the Caremark directors had taken active steps by adopting an ethical code, organizing a confidential reporting system, appointing a compliance officer and training employees, that was enough to preclude liability. *Id.* at 970.

248. *See* Enron Report, *supra* note 58, at 14 (explaining that "[h]igh risk accounting practices, extensive undisclosed off-the-books transactions, inappropriate conflict of interest transactions, and excessive compensation plans were known to and authorized by the board").

triad of available protections: exculpation, indemnification, and insurance.[249] Starting in the late 1980s, corporate codes began to permit charter provisions exculpating directors for negligent conduct.[250] Most public corporations have amended their charters to include these provisions.[251] In order to overcome exculpatory provisions, plaintiffs had to show the directors' bad faith in making the disclosure decision.[252] Moreover, the duty to disclose does not extend to corporations, so that if an exculpatory provision applies, shareholders cannot simply sue the corporation in lieu of the directors.[253]

As a result of the statutory and judicial litigation barriers outlined above, in the kinds of financial nondisclosures that appear to be at the root of many of the recent corporate defections, shareholders (who might wish to bring their claims against directors under the federal securities laws, for recklessly making a misleading statement in the Audit Committee Report,[254] for example) would have a difficult time meeting the PSLRA pleading standards. They would not be able to establish a state law fiduciary duty claim either because, to be liable, a director must knowingly misrepresent engaging in the described activities.[255]

---

249. *See, e.g.*, BAINBRIDGE, *supra* note 14, at 232 (discussing the triad of protections for directors).

250. *See* Mae Kuykendall, *Symmetry and Dissonance in Corporate Law: Perfecting the Exoneration of Directors, Corrupting Indemnification and Straining the Framework of Corporate Law*, 1998 COLUM. BUS. L. REV. 443, 467 (discussing the development of corporate codes permitting director indemnification).

251. *See* BAINBRIDGE, *supra* note 14, at 300 (discussing the evolution of state exculpatory statutes and their adoption by corporations). In addition to exculpatory provisions, "all states have statutory provisions authorizing director indemnification to some degree." *Id.* at 301. Thus, expenses for legal defense, and the advancement of those expenses, are widely available. *Id.* at 304.

252. *See* Zirn v. VLI Corp., 681 A.2d 1050, 1061-62 (Del. 1996) ("The record reveals that any misstatements or omissions that occurred were made in good faith."); *see also* Arnold v. Soc'y for Sav. Bancorp, Inc., 650 A.2d 1270 (Del. 1994) (concluding that although the proxy statements at issue contained an omission of material fact, the directors were immune from liability due to the corporation's exculpatory charter provision).

253. *See, e.g.*, Arnold v. Soc'y for Sav. Bancorp, Inc., 678 A.2d 533, 534 (Del. 1996) (explaining that shareholders seeking to hold the corporation liable for proxy statements containing materially misleading misstatements were precluded from suit because the federal proxy rules provided a remedy).

254. Item 306 of Regulation S-K and S-B and Item 7(d)(3) of Schedule 14A require firms listed on a national exchange to provide an Audit Report in the company's proxy statement disclosing whether the audit committee reviewed audited financial statements, discussed them with management, discussed matters with the independent auditor, and recommended that they be included in the annual report. *See* 17 C.F.R. §§ 229.306(a), 240.14a-101 (2001).

255. *See, e.g.*, Gregory S. Rowland, *Note, Earnings Management, the SEC, and Corporate Governance: Director Liability Arising from the Audit Committee Report*, 102 COLUM. L. REV. 168, 168, 196-97 (2002) (observing that the "directors would merely need to read the financial statements . . . engage in certain discussions regarding those statements . . . and obtain from (and discuss with) the outside auditor a statement of that outside auditor's independence" and concluding

III. BUREAUCRATIC ACCOUNTABILITY

An intriguing question is why, given that directors know that their investors need fundamental financial information, given that regulations make at least a minimum of disclosure about corporate finances mandatory, and given the market forces urging corporate signaling of trustworthiness to investors, corporate cover-ups occur. If investors want their firms to engage in some risky behavior in order to achieve profits, why cover it up? If the fear is that investors will walk away, why do it? Cover-ups may arise from incentives to act self-interestedly: there may simply be some directors who are "bad apples," strategically defecting and who—as game theory illustrates—need to be punished to reestablish cooperation in the investor/issuer game. Alternatively, directors who abdicate their duties to investors may have such strong incentives to act cooperatively with management that they are willing to forego their monitoring duties. On the other hand, directors may have convinced themselves that they were acting cooperatively (consciously acting in what they believed at the time to be in the best interests of the firm and its stakeholders), but were mistaken about their own strategic payoff or about the other players' strategies or payoffs.

In the Enron example, the directors at Enron could have made a real difference by refusing to approve financial statements and other disclosures, refusing to approve transactions that had no economic substance, refusing to waive provisions of their corporate code (such as the conflict of interest provisions) and ultimately, by firing officers. In their defense, the directors claimed that they had been misinformed by management.[256] Nonetheless, the Senate Subcommittee concluded that "overall the Board received substantial information about Enron's plans and activities and explicitly authorized or allowed many of the questionable Enron strategies, policies and transactions now subject to criticism."[257] No one has accused the Enron directors of directly lining their own pockets. It is still possible that the Enron board members were "bad apples," run amok with self-interest. For example, board members may have had a conflict of interest with respect to approving questionable accounting practices and disclosure failures because they were paid partly in stock options. Enron board members were compensated at about $350,000 per year (nearly twice the national average), in cash, restricted stock, phantom stock units (deferred cash

---

that "[t]oo many corporate managers, auditors, and analysts are participants in a game of nods and winks").

256. Throughout the Senate hearings, the directors who were interviewed maintained that management withheld key information from them. *See* Senate Permanent Subcommittee on Investigations Report, *The Role of The Board of Directors in Enron's Collapse*, SR 107-70, 107th Cong., 2d Sess., July 8, 2002 at 45.

257. *Id.* at 13, 14 ("High-risk accounting practices, extensive undisclosed off-the-books transactions, inappropriate conflict of interest transactions, and excessive compensation plans were known to and authorized by the Board.").

payments linked to the price of stock) and stock options.[258] The incentives to cooperate with management may have swamped the directors' incentives to monitor. They may have been reluctant for reputational or strategic reasons to do anything that would jeopardize their relationship with management.

In an earlier article, I argued that, given the information that they had, the directors may have convinced themselves that they were acting in the best interests of the corporation.[259] There, I argued that bounded rationality, the theory that human thinking evolved through repeated interactions with the environment, had consequences not only for what happened at Enron, but also for the success of the Sarbanes-Oxley Act's solution.[260] Heuristics and group interactions may affect corporate actions in predictably adverse ways.[261] I argued that understanding the operation of these heuristics in the organizational context and the conditions that make them more likely to occur may also illuminate ways to counter these tendencies.[262]

This Article extends the argument that these adaptive mechanisms, aiding quick and satisficing decisionmaking, may have effects on three aspects of the corporate climate that foster defections such as those at Enron. First, the overconfidence bias delineated by cognitive psychology as being especially prevalent in sales and marketing environments (and Enron at the time of its demise was primarily a derivatives trading entity) may have created a tendency to overrate the company's overall prospects, contributions and talents, making

---

258. Enron Report, *supra* note 58, at 11. Stock options are a worrisome source of pay because they enable the board member to benefit from stock gains without any risk of loss. *See id.* at 56.

259. *See* Erica Beecher-Monas, *Corporate Governance in the Wake of Enron: An Examination of the Audit Committee Solution to Corporate Fraud*, 55 ADMIN. L. REV. 357, 391 (arguing that "[g]roup dynamics under conditions of relative loss coupled with systemic cognitive biases that affect decisionmaking and behavior in a context of uncertainty help explain what happened at Enron, and also suggest the fallacy of the legislative solution").

260. *See generally* Simon; *see* David Liabson & Richard Zeckhauser, *Amos Tversky and the Ascent of Behavioral Economics*, 16 J. RISK & UNCERTAINTY 7 (1998) (surveying the literature on behavioral economics).

261. A broad spectrum of people under documented conditions, in particular contexts, make decisions that appear anomalous from a utility maximizing standpoint and may violate their own expressed preferences for decision making. *See* Robyn M. Dawes, *Behavioral Decision Making and Judgment, in* 1 HANDBOOK OF SOC. PSYCHOL. 498 (Daniel T. Gilbert et al., eds., 4th ed. 1998) (noting that these anomalies in rationality are systemic rather than ad hoc and are highly replicable in experimental settings). Far from being a pessimistic assessment of human rationality, however, these studies indicate contexts in which decision makers may need a structured process to make optimal decisions. *See, e.g., id.* at 500 (reporting studies showing that when anomalies in reasoning are "made transparent" to decisionmakers, their judgment improves); Bazerman at 8-10 (asserting that individual and group decision making can be improved by improved awareness of error-prone heuristics and by providing explicit strategies to counter them).

262. *See* Gerd Gigerenzer, *From Tools to Theories: A Heuristic of Discovery in Cognitive Psychology*, 98 PSYCHOL. REV. 254, 267 (1991) (frequency of anomalous reasoning can be decreased by making the probabilistic nature and questions explicit).

it willing to engage in novel ventures (such as the formation of the special purpose entities that ultimately destroyed the company) and left it with an inflated notion of its ability to control the risks its actions were creating.[263] Second, once embarked on the risky ventures, cognitive conservatism and the phenomenon of cognitive dissonance created a commitment bias that entrenched organizational commitment to the solution of special purpose entities and precluded remedial action.[264] Third, a strong bias in human decisionmaking to simplify (hence the need for heuristics and biases in the first place) is especially true of group decisionmaking, and when presented with complex financial transactions, directors have a tendency to think they are simpler and less controversial than warranted.[265] Although these cognitive quirks once may have provided an evolutionary advantage, they now impede wise corporate judgment, and any regulation ought to attempt to counteract these tendencies.

Congress's solution to the corporate debacles illustrated by Enron was the passage of the Sarbanes-Oxley Act. The Sarbanes-Oxley Act's principal remedy for director abdication of monitoring duties was to place the audit committee in charge of monitoring the financial controls of the corporation.[266] Because the legislation does not account for the way people interact in reaching their decisions, it is unlikely to prevent future incidents like Enron. Cognitive psychology and evolutionary game theory help explain why.

A. Corporate Compliance Programs: The Congressional Solution and Its Problems

Under the Sarbanes-Oxley Act, the board's audit committee must oversee corporate financial disclosures.[267] Although not all reporting companies are required to have an audit committee, if a firm wishes to be listed on a national exchange, it must have one.[268] The firm must also disclose whether the audit committee has a financial expert, and if not, explain why.[269] Under the new legislation, the audit committee not only has sole authority to appoint, compensate and oversee the firm's auditors, it must also oversee the firm's

---

263. Beecher-Monas, *supra* note 259, at 381-86.

264. *Id.* at 386-87.

265. *Id.* at 376.

266. *See* Sarbanes-Oxley Act § 301 (establishing the audit committee requirement for publicly listed companies); § 302 (requiring the signing officers to disclose deficiencies in internal controls to the audit committee).

267. Sarbanes-Oxley Act §§ 301, 302.

268. Section 301(1)(A) of the Sarbanes-Oxley Act requires the SEC to "direct the national securities exchanges and national securities associations to prohibit the listing of any security of an issuer that is not in compliance" with the audit committee requirements.

269. Sarbanes-Oxley Act of 2002, Pub. L. No. 107-204, § 407, 116 Stat. 745, 790 (2002). To qualify as a financial expert "the Commission shall consider whether a person has, through education and experience as a public accountant or auditor or a principal financial officer, comptroller, or principal accounting officer of an issuer" sufficient experience. § 407(b).

compliance program.[270] Moreover, section 302 of the Sarbanes-Oxley Act requires only that the principal executive and financial officers certify that they have disclosed deficiencies to the issuer's audit committee or persons fulfilling the equivalent function. Although the initial responsibility for setting up the program (and certifying that they have done so) rests on the corporation's CEO and CFO, they must report all "significant deficiencies" to the audit committee.[271] This effectively makes the audit committee the bearer of ultimate responsibility.

The idea of audit committee monitoring as a solution to financial reporting failures did not originate with Sarbanes-Oxley. The New York Stock Exchange and NASDAQ both required listed firms to have audit committees composed mostly or exclusively of independent directors since at least 1999.[272] Proxy statement disclosure about the audit committee's independence and discussions between the audit committee and management about audited financial statements have also been required since 1999.[273] However, neither the listing requirements nor the proxy statement disclosure rules were based on any empirical studies about audit committee effectiveness. Indeed, a number of studies found that the presence of an audit committee does not affect the likelihood of accounting fraud.[274]

Sarbanes-Oxley requires that the audit be composed of independent directors.[275] Independence is defined as meaning that the director may not accept consulting, advisory, or other fees or be an affiliated person of the issuer or its subsidiaries.[276] The Sarbanes-Oxley Act defines independence to more closely comport with an absence of conflict. The definition, therefore, may be a helpful clarification. However, the use of independent directors as a solution to the monitoring problem is no more empirically based than the audit committee solution. Most large firms already have a majority of independent directors, and

---

270. The audit committee must establish procedures for "the receipt, retention, and treatment of complaints received by the issuer regarding accounting, internal accounting controls, or auditing matters." Sarbanes-Oxley Act § 301(4)(A). In addition, the audit committee must resolve any disagreements between the auditor and management. § 301(2). Finally, because the CEO and CFO must report to the audit committee any deficiencies, fraud, or significant changes in the internal control system, under section 302, the audit committee appears to bear the ultimate responsibility for its oversight.

271. *See* Sarbanes-Oxley Act § 302 (requiring certification of the adequacy of the program).

272. *See* NYSE Listed Company Manual § 303.01(B)(2)(a); 17 C.F.R. §§ 229.306(a), 240.14a-101 (2001) (requiring listed firms to provide shareholders with an annual Audit Committee Report).

273. *See* SEC Release No. 34-42266 (Dec. 22, 1999).

274. Mark S. Beasley, *An Empirical Analysis of the Relation Between the Board of Director Composition and Financial Statement Fraud*, 71 ACCT. REV. 443 (Oct. 1996).

275. Not only must listed companies have an audit committee, but the Sarbanes-Oxley Act requires that if there is an audit committee, its members must be independent directors. Sarbanes-Oxley Act, Pub. L. No. 107-204, §301(3)(a), 116 Stat. 745, 775-77 (2002).

276. *Id.* § 301(3)(B)(1),(11).

audit committees were already expected to be independent.[277] Indeed, Enron appears to have had a model board since only two of its fourteen members were insiders.[278]

Like the audit committee solution to monitoring problems, the virtues of independent boards are debatable. Although the Business Roundtable recommended that a substantial majority of the directors be independent,[279] as did the National Association of Corporate Directors,[280] at least one study finds that the presence of inside directors on the board significantly increases the probability of accounting fraud,[281] even where the majority is composed of outside directors, insider presence on most boards remains a strong influence.[282] In addition, a number of studies show that firms with a majority of independent directors do not perform any better than firms without such boards, and that firms with only one or two inside directors may actually perform worse.[283]

Most of the audit committee's information will still come—directly or indirectly—through the CEO and CFO. Because the independent board members—if not the audit committee—will normally set the CEO's pay, and the CEO therefore has an incentive to paint a positive picture, there is reason to believe that the information reaching the audit committee may be skewed.[284] The audited financial reports ought to provide a check on this kind of misinformation,

---

277. *See* Sanjai Bhagat & Bernard Black, *The Uncertain Relationship Between Board Composition and Firm Performance*, 54 BUS. LAW. 921, 921 (1999) (citing a survey of 484 of the S&P 500 firms finding that over half had only one or two inside directors, the median firm and over 80% outside directors, and only nine firms had a majority of inside directors).

278. *See* Jeffrey N. Gordon, *What Enron Means for the Management and Control of the Modern Business Corporation: Some Initial Reflections*, 69 U. CHI. L. REV. 1233, 1241 (2002) (noting that the Enron board was "a splendid board on paper, fourteen members, only two insiders . . . . [T]he outsiders had relevant business experience . . . [and] owned stock . . . [but] was ineffectual in the most fundamental way").

279. THE BUS. ROUNDTABLE, STATEMENT ON CORP. GOVERNANCE 10 (1997).

280. Alan L. Dye, Securities Law Compliance Programs, ALI-ABA Course of Study, July 18-20, 2002.

281. Beasley, *supra* note 274, at 456-57 (using regression analysis of seventy-five fraud and seventy-five no-fraud firms to determine that no-fraud firms have a significantly higher proportion of "gray" and outside directors).

282. *See* Bhagat & Black, *supra* note 277, at 923 (noting that firms with majority inside directors perform equivalently to those with majority outside directors).

283. *See id.*; James P. Walsh & James K Seward, *On the Efficiency of Internal and External Corporate Control Mechanisms*, 15 ACAD. MGMT. REV. 421, 434 (1990) (citing studies showing that "there does not yet seem to be consensus support (either theoretically or empirically) for the conventional wisdom that either an increased presence of outsiders on the board of directors or the increased ownership stakes of any shareholder group (including management) necessarily improve corporate performance").

284. *See* Donald C. Langevoort, *The Human Nature of Corporate Boards: Law, Norms, and the Unintended Consequences of Independence and Accountability*, 89 GEO. L.J. 797, 812 (2001) (noting the "strong last-period temptation to manipulate the information given to the board").

but there are reasons to doubt their complete objectivity.[285]

The Sarbanes-Oxley Act's explicit focus on the audit committee as the ultimate firm overseer is new, however. This solution is troubling. Little evidence exists that compliance programs prevent misconduct. The deterrence model of regulation, based on the theory that legal compliance will occur whenever the pleasure and profits of breaking the law are outweighed by the pain of punishment, is problematic when applied to corporate behavior.[286] This is partly because of the diffusion of responsibility in corporations, and partly because laws regulating corporate conduct are not capable of defining the precise behavior required of corporate actors.[287] As a result, cooperative enforcement techniques, such as the federal sentencing guidelines'[288] (which permit the existence of an effective compliance program to reduce the severity of sentencing), and the Sarbanes-Oxley Act's requirement of internal controls monitored by the audit committee, have become a favored method of enforcement. The idea is that the corporation is in the best position to institute preventive and detection measures.[289] The goal is to provide incentives for optimal corporate behavior. An anticipated benefit is the reduction of public monitoring costs.

The idea underlying such regulation is a model of business regulation that

---

285. MAX H. BAZERMAN, JUDGMENT IN MANAGERIAL DECISION MAKING 2 (3d ed. 1994) (discussing the Phar-Mor audit to explain why auditors have a pervasive and intractable conflict).

286. *See* Michael A. Perino, *Enron's Legislative Aftermath: Some Reflections on the Deterrence Aspects of the Sarbanes-Oxley Act*, 76 ST. JOHN'S L. REV. 671, 674-76 (discussing the economic theory of deterrence); John T. Scholz, 60 LAW & CONTEMP. PROBS. 253, 258 (1997) (stating that the "simple deterrence model is most appropriate when legal statutes unambiguously define corporate misbehavior" but noting that "rules are seldom capable of defining the exact behavior desired of corporations").

287. Scholz, *supra* note 100, at 258 (observing that while "the simple deterrence model is most appropriate when the legal statutes unambiguously define corporate misbehavior" such rules are rare).

288. UNITED STATES SENTENCING GUIDELINES § 8C2.5(f) (reducing base fines in the presence of an effective compliance program).

289. *See* Jennifer Arlen & Reinier Kraakman, *Controlling Corporate Misconduct: An Analysis of Corporate Liability Regimes*, 72 N.Y.U. L. REV. 687, 693 (1997) (noting that "entity liability can lead companies to institute 'preventive measures' that deter by making misconduct more difficult or expensive for wrongdoers, or by reducing the illicit benefits of unpunished (or successful) misconduct, without affecting the probability that it is detected by enforcement officials"). In discussing the financial disclosure regime, the agents at issue are the managers, such as those at Enron who advocated the use of special purpose entities to keep debt off the books, the directors who waived conflict of interest proscriptions in the Enron compliance code, and both managers and directors who failed to disclose these shenanigans. The problem with entity liability for nondisclosure, however, is that it effectively permits the shareholders to be hit twice: once by the managers and directors who failed to disclose, and once by the imposition of liability on the corporation, a liability they must ultimately pay. Moreover, in an Enron-type situation, where the corporation itself is bankrupt, no entity remains for liability.

Ayres and Braithwaite call an "enforcement pyramid" strategy.[290] In this strategy, regulators predominantly rely on self-regulation (the base of the pyramid) but may become increasingly punitive in the face of noncompliance, with the ultimate threat of a "big gun" at the top of the pyramid for those firms that persist in noncompliance.[291] This model is based on the tit-for-tat strategy of evolutionary game theory.[292] It increasingly characterizes enforcement practice.[293]

Even if there were evidence that compliance programs were effective in deterring misconduct, there remains the question of why the audit committee should be singled out as the responsible actor. If the justification is that these independent financial experts are the only ones in the firm that can understand the complex corporate financing decisions faced by today's firms, that justification has little merit. Financial complexity impairs the ability of everyone, including management, to determine value.[294] Complexity should alert directors to a problem. If the directors do not understand what is going on, neither the investors (or analysts) are unlikely to have any deeper insight, and disclosure is not going to perform its function of revealing the value of the firm.

Corporate compliance programs ("internal controls") are supposed to solve the problem of ensuring that those at the top of the bureaucracy are informed regarding what is going on "down in the trenches." Like the audit committee and independent director solutions, the internal controls solution has been around for

---

290. *See* AYRES & BRAITHWAITE, *supra* note 100, at 39 (describing the enforcement pyramid).

291. *Id.* at 48-49 (stating that "regulators should not do without an image of invincibility in the background, and should be reluctant to push punishment to the foreground of day-to-day regulatory encounters").

292. *See* Sidney A. Shapiro & Randy S. Rabinowitz, *Punishment versus Cooperation in Regulatory Enforcement: A Case Study of OSHA*, 49 ADMIN. L. REV. 713, 727 (1997) (explaining how a tit-for-tat strategy combined with a range of sanctions, as with an enforcement pyramid, "can increase the potential of this strategy").

293. *See* Darryl K. Brown, *Street Crime, Corporate Crime, and the Contingency of Criminal Liability*, 149 U. PA. L. REV. 1295, 1312 (2001) (explaining that the model of an enforcement pyramid is "now widely accepted in regulatory debate and increasingly characterize[s] enforcement practice").

294. *See* Douglas G. Baird & Robert K. Rasmussen, *Four (or Five) Easy Lessons From Enron*, 55 VAND. L. REV. 1787, 1802 (2002) (explaining that "Enron's basic business plan—combining contracting over commodities with supplying the physical asset itself—created a large network of interrelated entities" which, although they had tax and accounting advantages, made it "difficult for those in charge to assess exactly how any given Enron division was performing"). As the authors note:

> One of the worst things a decisionmaker can do is pollute her own sources of information. The sheer complexity of understanding what Enron did and did not own undermined the business model premised upon the idea that a firm that combines the trading function with the delivery function enjoys a comparative advantage.

Id. at 1802-03.

a while.[295] Since the passage of the Foreign Corrupt Practices Act in 1978, the SEC has required registered or reporting companies to have a system of internal financial controls, and has imposed liability for failing to adequately maintain financial controls.[296] In the 1980s, the securities industry began creating internal compliance programs and routinely began to engage outside counsel for internal investigations.[297] In addition, recognizing the centrality of the problem of information and accountability to modern business, the Federal Sentencing Guidelines permit a reduction in criminal sentence for guilty corporations that have in place an effective corporate compliance program to monitor and assure the flow of information and prescribe steps to be taken in the event of a misstep.[298] Thus, the duty of care imposed by the guidelines requires firm managers to maintain adequate oversight of the firm's operations and to obtain adequate and reliable information before making decisions and taking action.[299]

Although such programs are already widespread,[300] the effectiveness of corporate compliance programs has been hotly debated. Since 1997, for example, despite the widespread adoption of compliance programs, there has been a significant growth in financial restatements to correct material

---

295. For a brief history of such legislative attempts to stimulate corporate self-policing, see Note, *The Good, the Bad, and Their Corporate Codes of Ethics: Enron, Sarbanes-Oxley, and the Problems with Legislating Good Behavior*, 116 HARV. L. REV. 2123, 2124 (2003).

296. *See* SEC v. World-Wide Coin Invs., Ltd., 567 F. Supp. 724, 749-50 (N.D. Ga. 1983) (discussing Securities Exchange Act § 13); *see also* COX ET AL., *supra* note 175, at 705 (discussing the legislative history of § 13(b)(2) and explaining that although a firm's materially misleading financial statements were actionable even before the Foreign Corrupt Practices Act, their passage clarified situations "where . . . the law governing the primary disclosure duty is quite fuzzy," such as where the deficiency is quantitatively immaterial but raises questions about "character, competence or integrity of management"). Courts have not permitted private causes of action to proceed under the accounting controls provisions. *See, e.g.*, Lewis v. Sporck, 612 F. Supp. 1316, 1332 (N.D. Cal. 1985).

297. Donald B. Ayer & James J. Graham, *Corporate Disclosure Programs: Voluntary and Mandatory*, 1 BUS. CRIMES BULL.: COMPLIANCE & LITIG. 4 (May 1994).

298. The Federal Sentencing Guidelines for Organizations reduce fines for violations that have taken place despite the presence of an effective compliance program. U.S.S.G. § 8C2.5(f), 8C2.6. This is based on the theory that organizational due diligence is a proxy for intent. I do not disagree with that theory. Rather, I disagree with the notion of a corporate partnership in crime control.

299. U.S.S.G. § 8A1.2, comment 3(k)(2); *In re* W.R.Grace & Co. Report, SEC Rel. No. 34-39157 (Sept. 30, 1997) (concluding that officers and directors of W.R. Grace & Co. failed to fulfill their obligations under the securities laws by failing to inquire into the reasons for nondisclosure in periodic reports of material information of which they were aware). At least one state law decision has weighed in (at least in dicta) opining that failure to have such a program would be a breach of state law fiduciary duty. *In re* Caremark Int'l Inc. Derivative Litig., 698 A.2d 959 (Del. Ch. 1996).

300. *See* Stephen Calkins, *Corporate Compliance and the Antitrust Agencies' Bi-Modal Penalties*, 60 LAW & CONTEMP. PROBS. 127, 147 n.84 (1997) (citing a survey showing that 75 to 95% of U.S. firms have written codes of conduct).

misrepresentations of financial results, primarily due to revenue recognition issues.[301] While many commentators extol the ability of compliance programs to ensure the flow of information within the firm,[302] there are a number of valid criticisms of such programs. One criticism is that the only effect of such programs has been to shift the locus of liability further down the corporate hierarchy.[303] People who run the company will attempt to minimize their own risks. As a consequence, the senior personnel have incentives to report fraud only if there is little chance that they will be implicated.[304]

Because these programs create many cosmetic rather than real changes, legislators have attempted to give them some force.[305] An effective compliance

---

301. United States General Accounting Office, Report on Financial Statement Restatements (Oct. 2002), at 14-15 (noting that while the number of listed companies decreased by 20% from 1997-2002, the number of listed companies restating their financials increased by 165% to a projected 3% of listed companies by the end of 2002). Over this period the average (median) market capitalization of a restating company grew from $500 million ($143 million) in 1997 to $2 billion ($353 million) in 2002. *Id.* at 17. The GAO database excluded announcements of restatements for reasons other than material misstatements of financial results. *Id.* at 21. For example, the WorldCom restatements involved overstating net income by recording operating expenses as capital expenditures. *See* WorldCom Press Release, WorldCom Announces Intention to Restate 2001 and First Quarter 2002 Financial Statements (June 25, 2002).

302. *See, e.g.*, Ribstein, *supra* note 2, at 15 (opining that "[e]stablishment of control systems within the firm and protecting whistleblowers helps ensure the flow of information within the company"); Lynn L. Dallas, *A Preliminary Inquiry into the Responsibility of Corporations and Their Directors for Corporate Climate: The Psychology of Enron's Demise*, ST. JOHN'S L. REV. (forthcoming) (proposing increased implementation of ethical compliance programs); Diana E. Murphy, *The Federal Sentencing Guidelines for Organizations: A Decade of Promoting Compliance and Ethics*, 87 IOWA L. REV. 697 (2002) (discussing the positive effects of the corporate compliance programs induced by the Federal Sentencing Guidelines); H. Lowell Brown, *The Corporate Director's Compliance Oversight Responsibility in the Post-Caremark Era*, 26 DEL. J. CORP. L. 1 (2001) (advocating increased use of internal control systems); Pamela H. Bucy, *Corporate Ethos: A Standard for Imposing Corporate Criminal Liability*, 75 MINN. L. REV. 1095, 1099 (1991) (published shortly before the organizational guidelines became effective; advocating the implementation of compliance programs).

303. *See* William S. Laufer, *Corporate Prosecution, Cooperation, and the Trading of Favors*, 87 IOWA L. REV. 643, 648-49 (2002) (discussing the trend toward "reverse whistleblowing" in corporations as a result of the Organizational Sentencing Guidelines' incentives structure); William S. Laufer, *Corporate Liability, Risk Shifting, and the Paradox of Compliance*, 52 VAND. L. REV. 1343, 1350 (1999) (opining that firms "have been extraordinarily successful in shifting both the locus of liability risk and the enforcement function down the corporate hierarchy").

304. *See* James J. Graham & Morris Silverstein, *Voluntary Disclosure Bandwagon: Pitfalls for Federal Agencies*, 2 BUS. CRIMES BULL.: COMPLIANCE & LITIG. 2, 2 (1995) (observing that in the defense industry, although voluntary disclosure programs resulted in increased reporting of minor accounting discrepancies, few of the more serious crimes involving high level personnel were uncovered).

305. *See, e.g.* Note, *supra* note 295, at 2127 (noting critics who assert that compliance codes

program risks generating incriminating information, making truly effective programs unduly expensive to the corporation's management and thereby decreasing incentives to engage in meaningful self-evaluation.[306] Another critique of the purported partnership in fighting crime is that the pervasiveness of agency costs makes optimal compliance unlikely.[307] Moreover, the risks associated with internal investigations—the backbone of any accountability program—may create incentives for half-hearted compliance. There is a risk that the information generated by an internal investigation may identify a problem that will force affirmative action to avoid liability; this risk is a deterrent to serious investigatory efforts.[308] Also present is the threat that the investigation may uncover discoverable documents that could be used against the firm and its managers in private civil litigation.[309] Further, a firm that does not properly implement and enforce a rigorous compliance program may be in worse shape—by exposing itself to grater liability—than a firm with a less rigorous code.[310]

Compliance programs are not cheap. They are costly to set up and run, and the people employed in the compliance office will have their own rent-seeking agendas.[311] Not only is disclosure involving highly placed personnel rare, but supervisory personnel are in the best position to frame any disclosure and thus shape the facts. They are also expensive in the sense that it is difficult to tell how effectively the monitoring is being done.[312] The theoretical justification for

---

"comprise little more than platitudes" and asserting "the need for corporate codes with teeth"). Both codes of ethics and internal controls programs are part of corporate compliance, so although the Note addresses § 406 (the code of ethics requirement) rather than the § 404 internal controls requirement, the concerns are similar for both. *Id.* at 2138 (noting the problem).

306. *See,* Jay P. Kesan, *Encouraging Firms to Police Themselves: Strategic Prescriptions to Promote Corporate Self-Auditing*, 2000 U. ILL. L. REV. 155 (discussing the perils of self-evaluation with respect to environmental laws).

307. *See* Donald C. Langevoort, *Monitoring: The Behavioral Economics of Corporate Compliance with Law,* 2002 COLUM. BUS. L. REV. 71, 111 (2002). This concern for conflicting interests is addressed in two major changes wrought by the Sarbanes-Oxley Act, one prohibiting auditors from providing contemporaneous consulting services to the firms they audit (§ 204(a)), and one prohibiting firms from extending personal loans (or arranging them) to officers and directors (§ 402(a)).

308. Michael P. Kenny & William R. Mitchelson, Jr., *Corporate Benefits of Properly Conducted Internal Investigations*, 11 GA. ST. U. L. REV. 657, 661 (1995).

309. *Id.*

310. *See, e.g.*, Harvey L. Pitt & Karl A. Groskaufmanis, *Minimizing Corporate Civil and Criminal Liability: A Second Look at Corporate Codes of Conduct*, 78 GEO. L.J. 1559, 1560 n.8 (1990) (noting the dangers of implementation).

311. *See* Langevoort, *supra* note 284, at 100 (noting that "[a]ll economic units within an organization tend to construe ambiguous information in a self-serving way that maximizes its influence and, hence, claim to additional resources").

312. Arlen and Kraakman argue that the optimal level of deterrence can be provided by a mixture of high penalties and a self-policing duty, although they acknowledge the difficulties of

imposing the costs of internal investigations is that the organization is in the best position to detect and deter aberrant behavior. These costs, however, are not imposed upon the people who can detect and deter misconduct, but on the shareholders, who have little power to affect management decisions.[313]

B. The Risky Shift and Strategic Interactions

Corporate compliance programs, like the system of internal financial controls mandated by Sarbanes-Oxley, although praised as a partnership in crime control and meant to align the interests of business with the public interest,[314] are ineffective in deterring corporate misconduct, because they are not based on the decisionmaking processes of interacting groups.[315] The context where monitoring counts is where the firm is experiencing conditions of relative loss. Although classical economists claimed that people are naturally risk averse, this is not necessarily so. Instead, people appear to avoid risky actions only when they are experiencing relative wealth.[316] They favor risky actions when they are in a losing situation.[317] For example, when all options are undesirable, high risk gambles are often preferred to fairly certain losses.[318] Daniel Kahnemann and Amos Tversky proposed a formal model that they called prospect theory, in which people are both risk-seeking and risk-averse: risk-averse for moderate probabilities and risk-seeking for small probabilities of gain; the opposite for probabilities of loss.[319] When this kind of relative loss occurs, cover-ups are

---

proving optimal self-policing. *See* Arlen & Kraakman, *supra* note 289, at 716 (noting that "a duty-based regime can solve the credibility problem only if the court can determine whether the firm has implemented efficient enforcement measures").

313. *See, e.g.*, Jonathan R. Macey, *Agency Theory and the Criminal Liability of Organizations*, 71 B.U. L. REV. 315, 327 (1991) (observing that under conditions of near-bankruptcy, there are strong managerial incentives for misconduct in order to save their jobs, and that when there has been a period of lax enforcement, corporate culture may permit adoption of profitable but illegal practices).

314. *See* JAY A. SIGLER & JOSEPH E. MURPHY, INTERACTIVE CORPORATE COMPLIANCE: AN ALTERNATIVE TO REGULATORY COMPULSION (1988) (describing the growth of the idea of government-business cooperation).

315. Beecher-Monas, *supra* note 259, at 357 (describing cognitive biases undercutting the effectiveness of compliance programs).

316. *See* Amos Tversky & D. Kahnemann, *Rational Choice and the Framing of Decisions*, 59 J. BUS. 251-94 (1986) (studies showing that people treat risks concerning perceived gains differently from risks concerning perceived losses).

317. *See* T. S. Bateman & C.T. Zeithaml, *The Psychological Context of Strategic Decisions: A Model and Convergent Findings*, 10 J. STRATEGIC MGMT. 59-74 (1989).

318. *See* Daniel Kahneman & Amos Tversky, *Prospect Theory: An Analysis of Decision Under Risk*, 47 ECONOMETRICA 263-90 (1979).

319. *See generally id.* at 263-91.Since its original proposal, prospect theory has been tested extensively. *See, e.g.*, George Wu et al., *Decision Under Risk*, *in* BLACKWELL HANDBOOK OF JUDGMENT AND DECISION MAKING (Nigel Harvey & Derek Koehler, eds. forthcoming) (reviewing

predictable.[320]

Enron began to experience intense financial pressures immediately before the board approved the risky—and conflicted—structured financing transactions that ultimately led to the firm's demise. Although Enron appeared highly profitable between 1995-2000, when its revenue grew at a compound annual rate of more than 60%, its investments in energy and water plants and fiber optic networks were sizeable and did not produce the anticipated income stream.[321] After 1997, Enron's cost of capital consistently exceeded its return on invested capital, and its annual return on invested capital decreased from October 1995 until Enron finally declared bankruptcy in December 2001.[322] The transformation of Enron's old-fashioned energy business—oil and gas pipelines, power plants, etc.—into an online energy trading business, similar to a derivatives trading company, made

---

the literature and discussing the original prospect theory and its later refinements, evaluating their strengths and weaknesses).

320. This risky shift presents the familiar "last period problem," in which, as Arlen and Kraakman discovered, most open market securities frauds are prompted by the self-interested fears of management that adverse financial results—periods of relative loss—will cause them to lose their jobs. *See* Arlen & Kraakman, *supra* note 289, at 724-30 (demonstrating that most securities frauds occur when firms face the threat of insolvency, and arguing that it is rational for managers to postpone or avoid disaster by taking escalating risks that culminate in corporate cover-ups, including securities fraud). Management takes risks to buy time during which a corporate turn around may be possible, or to hang on as long as they can. *See* Donald C. Langevoort, *Capping Damages for Open-Market Securities Fraud*, 38 ARIZ. L. REV. 639, 655 (1996) (discussing the findings of both Jennifer Arlen and William Carney). When businesses experience financial reverses, agency incentives tend to reinforce the risk preference bias predicted by prospect theory. *See* Richard W. Painter, *Lawyers' Rules, Auditors' Rules, and the Psychology of Concealment*, 84 MINN. L. REV. 1399, 1419 (2000); Dan J. Laughmunn et al., *Risk Preferences for Below Target Returns*, 26 J. MGMT. SCI. 1238 (1980) (documenting managerial risk seeking under conditions of relative loss). Thus, corporate cover-ups are predictable if the corporation faces bankruptcy. *See* Arlen & Kraakman, *supra* note 289, at 724-27 (discussing the predictability of corporate nondisclosure in periods where the corporation faces insolvency). Corporate governance innovations devoted to reduce managerial risk aversion and encourage a more entrepreneurial risk-taking perspective (such as compensation that rewards short-term stock price gains) may actually exacerbate a firm's downward spiral. *See, e.g.*, Leo E. Strine, Jr., *Derivative Impact? Some Early Reflections on the Corporation Law Implications of the Enron Debacle*, 57 BUS. LAW. 1371, 1374 (2002). The author noted that over

> the last two decades, much thought has been devoted to finding ways to direct the attention of boards and directors away from a safe managerialist perspective focusing on entity preservation, and toward a more entrepreneurial, risk-taking, and competitive-enhancing attitude . . . [such as] the implementation of compensation policies for managers and directors that reward short-term stock price gains.

321. *See* Stuart L. Gillan & John D. Martin, Working Paper Series, *Financial Engineering, Corporate Governance, and the Collapse of Enron*, U. Del. Center for Corporate Governance 2002, at 7 (2002), http://www.be.udel.edu/ccg/research.htm (last visited Oct. 22, 2003).

322. *Id.* at 37.

significant cash flow imperative.[323] Because it needed significant amounts of daily cash to settle its contracts, Enron depended upon large lines of credit, the availability of which were based on Enron's credit ratings.[324] Credit ratings in turn are based on a firm's financial statements. Therefore, to obtain investment-grade credit ratings, Enron had to increase cash flow, lower debt, and prevent large earnings fluctuations.[325] This loss-framing may have influenced the subsequent events. In order to accomplish these goals, the company decided on a strategy of shedding (or increasing immediate returns on) company assets like power plants that had low returns and persistent debt.[326] Because Enron was not able to find any buyers for these assets, these financial pressures created an intensifying debt burden.[327] This is precisely the kind of relative loss that predicts risky behavior.

These pressures, framed in terms of relative losses, created a context for taking risks that placed the firm in great danger and ultimately caused its demise. Enron reported more gain on its operations than it made.[328] Because it could not find buyers for its assets, it sold them to "unconsolidated affiliates."[329] It engaged in many transactions that gained it nothing other than an ability to hide its finances.[330] It exposed itself to contingent liabilities through its special purpose entities.[331] It also permitted these off-balance sheet special purpose entities to be run by Enron officers, a conflict prohibited by Enron's compliance program. These decisions were made by (or under the direction of) Enron's directors, who operated under a state-of-the-art corporate governance structure. In hindsight, those were terrible decisions. The important question now, however, is whether Sarbanes-Oxley can prevent future debacles. Optimism and the illusion of control tend to increase risk taking.[332] This is not a bad thing under normal conditions, when managers and directors tend to be risk averse and

---

323. *See* Enron Report, *supra* note 58, at 7 (noting that Enron's online energy trading business bought and sold contracts to deliver natural gas, oil or electricity, treating them like commodity futures—but outside the regulatory purview of the securities or commodities laws).

324. *Id.*

325. *Id.*

326. As the presentation made to the Finance Committee explains, "[l]imited cash flow to service additional debt," combined with "[l]imited earnings to cover dilution of additional equity" meant that "Enron must syndicate . . . in order to grow." *Id.* (quoting Finance Committee Presentation (Oct. 2000)).

327. *Id.*

328. Baird & Rasmussen, *supra* note 294, at 1801.

329. *See* Enron Report, *supra* note 58, at 7-8 (noting the Enron board's "intense focus on its credit rating, cash flow, and debt burden").

330. *See* Baird & Rasmussen, *supra* 294, at 1804 (citing the example of the Raptor III transaction, for which it paid LJM2 $39.5 million without gaining anything "other than an ability to hide its finances from investors for losses over the short term").

331. Gillan & Martin, *supra* note 321, at 9.

332. Daniel Kahnemann & Dan Lovello, *Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking*, 39 MGMT. SCI. 17, 28 (1993).

loss averse.[333] But when the company is experiencing financial reversals—under conditions of loss—where they are apt to be risk preferring, it could spell disaster. For example, Enron's management apparently persuaded itself—despite massive evidence to the contrary—that what worked in the context of deregulating energy markets in the United States could work in other markets globally.[334] As a result of this overconfidence, their financial models appear to have been hopelessly inaccurate.

C. The Role of the Board of Directors

In many ways, the Enron board appears to have met or exceeded the standards set by Sarbanes-Oxley, at least in form. For example, Enron's board had both an audit committee and a finance committee.[335] The audit committee advised the board on hiring the firm's independent auditor, ensured that the auditor was accountable, reviewed the auditor's compensation, reviewed the firm's annual financial statements, reviewed the financial statements included in the Annual Report to Shareholders, footnotes and management commentaries, and Form 10-K fillings, approved major changes and other choices regarding the appropriate accounting principles and practices to be followed in preparing the financial statements, assessed the firm's internal financial control systems, approved for recommendation to the board the corporate compliance policies and procedures, and filed a report in the annual proxy statement.[336] The finance committee was responsible, among other things, for monitoring management financial policy, plans and proposals, changes in risk management policy and the transaction approval process.[337]

Despite having what appeared to be an ideal corporate governance structure, Enron's board and subcommittees made a number of decisions that accelerated its shift toward risky alternatives.[338] First, in February 1999, the audit committee made the decision to reappoint as auditor Arthur Anderson, a questionable decision in light of the conflict between Anderson's role as a financial consultant and its role as auditor of the same transactions it had recommended.

---

333. As Kahneman & Lovallo note, "Bold forecasts and timid attitudes to risk tend to have opposite effects." *Id.* at 30. Countering risk aversion may be one of the reasons for prizing managerial optimism. However, under conditions of risk preference, overconfidence could lead to very bad judgment. *See id.* (noting that "[i]ncreasing risk taking could easily go too far in the presence of optimistic forecasts").

334. *See* Baird & Rasmussen, *supra* note 294, at 1797.

335. ENRON CORP., AUDIT AND COMPLIANCE COMMITTEE CHARTER (as amended Feb. 12, 2001), *in* ENRON CORP. PROXY STATEMENT PURSUANT TO SECTION 14(A) OF THE SECURITIES AND EXCHANGE ACT OF 1934, at 44-47 (2001). It also had compliance, compensation and management development, executive, nominating, and corporate governance committees. *Id.*

336. *Id.*

337. *Id.*

338. Enron Report, *supra* note 58, at 12 (observing that "more than a dozen incidents over 3 years . . . should have raised Board concerns").

In addition, Anderson had acknowledged that Enron's accounting practices (practices Anderson had consulted on) were highly risky.[339] This decision was revisited, with similar results, once or twice a year thereafter.[340] A second series of decisions entailed the waiver of conflicts of interest provisions in Enron's code. At a special meeting of the board in June 1999 (held by teleconference), the entire board approved the formation of a special purpose entity, LJM, to be owned and managed by CFO Fastow, without the prescribed approval of the finance committee.[341] This decision entailed waiving the conflict of interest provisions to permit CFO Fastow to manage the special purpose entity, but the board failed to put any controls in place to monitor this self-dealing transaction.[342]

Twice more in 1999 and 2000 the board approved similar special purpose entities allowing the firm's CFO to set up special purpose entities to improve Enron's financial statements.[343] Not until February 2001 did the board (through the compensation committee) request any information about the extent to which the CFO was personally benefitting from these transactions, and when the information was not forthcoming, the matter was dropped.[344] The board continued to approve structured financing transactions until by October 2000 nearly half of its assets were in Enron's "unconsolidated affiliates."[345] For many of these entities, the board also approved guaranteeing the off-book entity's debt.[346]

These facts may be unique to Enron.[347] The decisions of Enron's directors, however, illustrate a number of reasons why the Sarbanes-Oxley Act's reliance

---

339. *Id.* at 17 (citing Audit Committee Minutes).

340. *Id.* (citing Audit Committee presentations 1999-2001).

341. *Id.* at 24.

342. Although the board's ratification of the CEO's waiver of the firm's code of conduct was not required, it was explicitly requested at each of the LJM presentations. *Id.* at 24 (citing board presentations).

343. *Id.* at 12.

344. *Id.* at 32-33 (noting that it was not until a *Wall Street Journal* article appeared on October 19, 2001 that the board decided to place Fastow on leave) (citing *Enron CFO's Partnership Had Millions in Profit*, WALL ST. J., Oct. 19, 2001).

345. *Id.* at 8 (citing October 2000 presentation to the finance committee).

346. *Id.* at 12 (citing the example of Whitewing, where the board approved moving an affiliated company off Enron's books while guaranteeing its debt with Enron stock, and noting that "Committee and Board presentations throughout 1999, 2000, and 2001 chronicled the company's foray into more and more off-the-books activity.").

347. As Professor Coffee notes, "the problem with viewing Enron as an indiction of any systematic governance failure is that its core facts are maddeningly unique." Coffee, *supra* note 2, at 1403 (opining that "the passive performance of Enron's board of directors cannot fairly be extrapolated and applied as an assessment of all boards generally"). Although the particular decisions the Enron board made were unique and the product of the innovative circumstances of Enron, the decisions illustrate group dynamics under conditions of loss that are common to all large publicly held corporations.

on corporate compliance programs to achieve accountability is misconceived. First, a large percentage of directors of publicly held corporations are CEO's of their own firms.[348] Directors who manage their own firms and who have themselves been advocating structured financing as a way of improving financial appearances[349] may be inclined to see them as legitimate when asked to approve them as directors. CEO's who are reluctant to disclose facets of the internal workings of their own firms may be predisposed to tolerate iffy disclosure as directors. People who, as CEO's, have urged the appointment of auditors who consulted on the structured financing they were to audit may be less inclined to object when they are asked, as directors, to approve such appointments.[350]

In addition, the audit committee—a group of financial experts—may be more prone to overconfidence than the board as a whole,[351] making their placement as monitors-in-chief counterproductive. Moreover, confidence tends to be highest when people believe that there is consensus for their opinion and that a decision must be made quickly.[352] This is all too often the context of board decisions.

The unconscious heuristics and biases discussed above explain only part of the decisionmaking process. Strategic interactions also come into play. Reciprocity pays, especially in repeat interactions (such as those in the management/director game) within a small group, where people can keep track of previous behavior.[353] Although outside directors have their own corporate

---

348. *See* Ronald J. Gilson & Reinier Kraakman, *Reinventing the Outside Director: An Agenda for Institutional Investors*, 43 STAN. L. REV. 863, 874-75 (1991) (noting that "63 percent of outside directors of public companies are chief executive officers of other public companies").

349. Structured financing is prevalent among large firms, although most do not push the envelope as strenuously as Enron appears to have done. *See* Schwarcz, *supra* note 75, at 1309 (examining the differences between "the trillions of dollars of supposedly 'legitimate' securitization" and that of Enron).

350. *See* Paul E. Jones & Peter H. Roelofsma, *The Potential for Social Contextual and Group Biases in Team Decision-Making: Biases, Conditions and Psychological Mechanisms*, 43 ERGONOMICS 1129, 1144 (2000) (noting that one explanation of group polarization dynamics is that it "stems from people's motivation to be perceived, and to present themselves, in a favourable light").

351. *See* Mark Seidenfeld, *Cognitive Loafing, Social Conformity, and Judicial Review of Agency Rulemaking*, 87 CORNELL L. REV. 486, 498 (2002) (noting that overconfidence in their predictions is a "bias to which experts may be more prone than novices"); *see also* Philip E. Tetlock et al., *Assessing Political Group Dynamics: A Test of the Groupthink Model*, 63 J. PERSONALITY & SOC. PSYCHOL. 403, 419 (1992) (testing the empirical basis and theoretical logic of the groupthink model and noting that "groupthink promoted rigid and self-righteous patterns of thinking").

352. Ruben Orive, *Group Consensus, Action Immediacy, and Opinion Confidence*, 14 PERSONALITY & SOC. PSYCHOL. BULL. 573 (1988) (studies). The conditions of consensus and immediacy appear to have been present in each of the troubling Enron decisions. For example, the Enron board reported nearly unanimous decisions that were typically made within one to two hours. *See* Enron Report, *supra* note 58, at 8.

353. *See* VERNON L. SMITH, BARGAINING AND MARKET BEHAVIOR 117 (2000) (explaining that

cultures (since they are drawn primarily from executive positions in other firms, academia and government[354]), they nonetheless interact over time (at least five of the Enron board members, for example, had been with the firm since 1985, and the average in 2000 was eight years of service[355]).

Thus, their incentives for acting cooperatively toward each other may have outweighed the incentives to act cooperatively toward their investors. When people have to decide between competing courses of action, it is much easier to discount harm to statistical others—the investors, for example—than to people you know.[356]

In the corporate governance game played by managers and directors, the players' strategies is a repeated prisoners dilemma game, in which the players must decide whether to observe their fiduciary duties or evade them. Management's options are to comply with their fiduciary duties or evade them, while directors' options are to adopt a cooperative or deterrent enforcement strategy. If both cooperate, that is the social optimum because directors will not have to expend resources (time or energy) in monitoring and management can concentrate on meeting business goals rather than worrying about minor infractions. If directors behave cooperatively, however, management has incentives to behave opportunistically and if directors know that management will comply with its fiduciary duties, it may have incentives to nit-pick the details of management proposals rather than concentrating on business goals. The temptations to defect may cause directors to undertake an inefficient amount of scrutiny, while management may forego innovative business opportunities. A tit-for-tat strategy (or its kin, a firm-but-fair strategy) may avoid this inefficient outcome by having directors adopt a more discretionary monitoring strategy during periods where management's temptation to defect is lower (periods of relative gain, for example) or where there is little danger of conflict of interest. Joint cooperation or joint defection are the only plausible equilibria.[357]

---

reciprocity, in which "individuals incur short-term costs for their sharing in exchange for delayed benefits for others' sharing," is "possible in close-knit communities because each individual can 'keep score' and punish free-riders with sanctions" but "break down where sociability is pushed to the edge of credibility").

354. Sixty-three percent of outside directors are CEO's of other public companies. Gilson & Kraakman, *supra* note 348, at 874-75. As Gilson & Kraakman observe, "[they] are unlikely to monitor more energetically than they believe they should be monitored by their own boards." *Id.* at 875. Moreover, "personal and psychic ties to the individuals who are responsible for one's appointment" as well as "substantial director compensation" may further act to align directors' interests with management's rather than with the shareholders. Charles M. Elson, *Director Compensation and the Management-Captured Board—the History of a Symptom and a Cure*, 50 SMU L. REV. 127, 161-62 (1996).

355. Enron 2001 Proxy Statement.

356. *See* George Lowenstein, *Behavioral Decision Theory and Business Ethics: Skewed Trade-Offs Between Self and Other*, *in* CODES OF CONDUCT 214-15 (David M. Messick & Ann E. Tenbrunsel, eds. 1996) (discussing the trade-offs between options and courses of action).

357. For a discussion of the economics of the game-theoretic model, see Scholz, *supra* note

When the directors' sucker payoff (for management defection) exceeds its punishment payoff, however, capture by management is likely. Rather than leading to joint defection (i.e., stricter director scrutiny of management proposals and increased willingness to vote them down), when this happens, management defection may lead to the management defect: director cooperate equilibrium, a result that is not socially optimal. The social ties that directors frequently have with management, the identification of directors who are CEOs of their own firms with management, and the effort that must be expended in monitoring and uncovering information all make this capture equilibrium more likely.[358]

Regulators and the regulated community are also repeat players in this sense. Cooperation is likely to evolve in such interactions, but the interactions may become corruption and capture as well as optimal business behavior.[359] Agencies (such as the SEC) are much more likely to cooperate than defect, because monitoring and punishment are costly.[360] The theory of an enforcement pyramid is that legal compliance is more likely to be effective when a regulatory agency begins by cooperating with the regulated (here a corporation) in a tit-for-tat game, coaxing compliance through self-regulation, with a graduated series of sanctions for infractions, culminating with severe punishment for severe and repeated violations.[361] The Sarbanes-Oxley model of regulation mandates self-regulation in the form of the internal controls system and audit committee monitoring, and it has increased the severity of sanctions for violations.[362] The punishments are still graduated, because penalties will be enforced in conjunction with the Federal Sentencing Guidelines.[363] Thus, the model employed by Sarbanes-Oxley depends on self-regulation and graduated sanctions. What is missing, however, is the third face of Ayres' and Braithwaite's

---

100, at 188.

358. *See, e.g.*, Charles M. Elson & Christopher J. Gyves, *The Enron Failure and Corporate Governance Reform*, 38 WAKE FOREST L. REV. 855, 857 (observing that "board oversight may be doomed" because "directors who must monitor the managers have been appointed by the very managers they must monitor" and this creates "a great incentive for passivity and acquiescence to management's initiatives and little incentive to actively monitor").

359. *See* AYRES & BRAITHWAITE, *supra* note 100, at 55 (noting that the "conditions that foster the evolution of cooperation are also the conditions that promote the evolution of capture and indeed corruption").

360. *Id.* at 70 (noting that "firm defection must be of extraordinary proportions to overcome the attitudinal resistance of regulators to punishment" and suggesting tripartisim as a solution).

361. *Id.* at 35-40 (outlining the structure of the enforcement pyramid).

362. For example, earnings restatements due to material noncompliance will require the CEO and CFO to reimburse the corporation for any bonus received during the period covered by the restatement under 304; increased criminal penalties for securities violations under 807, 1106; increased criminal penalties for mail and wire fraud under 903 and ERISA under 904; making failure to certify financial reports a crime under 906.

363. *Cf.* Laufer, *supra* note 303, at 644 (noting that the Federal Sentencing Guidelines constitute "a dynamic enforcement game backed by a 'tall enforcement pyramid'").

enforcement pyramid, the participation of public interest groups.[364] Investor litigation might have functioned as public interest groups, enabling the enforcement pyramid.[365] These groups have been disabled by PSLRA and *Central Bank*, as well as judicial hostility to securities class actions, however.[366]

Arguably, increasing the size of the penalty, as Sarbanes-Oxley did,[367] might have an effect on deterring corporate misconduct. Under an economic notion of deterrence, either increased enforcement or increased penalties will deter misconduct.[368] Empirically, however, increasing the size of the penalties does not appear to have an effect on misconduct.[369] On the contrary, increasing the size of the penalty appears to be counter-productive, because it reduces monitoring.[370]

Sarbanes-Oxley also purports to increase SEC monitoring.[371] If believed,

---

364. *See* AYRES & BRAITHWAITE, *supra* note 100, at 56 (advocating tripartism by permitting public interest groups to "become a fully fledged third player in the game" that "can directly punish the firm" and thus "secure the advantages of the evolution of cooperation while averting the evolution of capture and corruption").

365. Although Ayres and Braithwaite conceived of the public interest group participation on the front end of the regulatory process, by providing them with full information about the deals cut between regulators and the regulated, they acknowledge that "back-end standing is a prerequisite for front-end submissions to be taken seriously." *Id.* at 77-78.

366. *See supra* Part II.E.2.b.

367. For example, maximum penalties for mail and wire fraud have been increased from five to twenty years, under Sarbanes-Oxley Act § 903; under § 1106 (amending 15 U.S.C. § 78ff(a)) securities fraud violations under the 1934 Securities Exchange Act have been increased for individuals to fines of $5 million (from $1 million) and terms of 20 years (from 10 years) and for organizations fines have been increased to $25 million from $2.5 million. These provisions are unlikely to actually change expected penalties, however, because maximum statutory sentences merely set an outside limit for the sentence. *See* Perino, *supra* note 286, at 686 (observing that the "penalty enhancements are unlikely to deter corporate crime to any greater degree than current provisions").

368. *See, e.g.*, George Tsebelis, *The Abuse of Probability in Political Analysis: The Robinson Crusoe Fallacy*, 83 AM. POLI. SCI. REV. 77, 79 (explaining and debunking the economic theory that the expected utility of misconduct depends on the size of the penalty).

369. *See* Scholz, *supra* note 100, at 255 (noting that although the level of compliance increases after penalties are imposed, in OSHA inspections "the size of the penalty has little impact on safety improvements, contradicting the basic premise of deterrence theory that large expected penalties explain compliance").

370. *See* George Tsebelis, *Penalty and Crime: Further Theoretical Considerations and Empiriccal Evidence*, 5 J. THEORETICAL POLITICS 349 (1993) ("[I]ncreases in penalty have no impact on crime, but reduce police monitoring.").

371. For example, maximum penalties for mail and wire fraud have been increased from five to twenty years under Sarbanes-Oxley § 903; under § 1106 (amending 15 U.S.C. § 78ff(a)), securities fraud violations have been increased for individuals (from $1 million to $5 million and from ten years to twenty years) and organizations (from $2.5 million to $25 million). These provisions are unlikely to actually change expected penalties, however, because prosecutorial

such statements about increased enforcement will save time, money and effort.[372] There are some reasons to be skeptical about the effectiveness of this measure. Even if such statements do result in increased monitoring and misconduct decreases as a result, the result will tend to be cyclical because as soon as misconduct decreases, so will enforcement.[373] Thus, "certainty and severity of penalty are inversely related."[374] It is not the size of the "big gun"[375] at the top of the enforcement pyramid that affects the rate of corporate misconduct but the rate of monitoring and enforcement.

IV. A NEW PROPOSAL: MONITORING THROUGH SELF-INSURANCE

Given the foibles of human decisionmaking under conditions of uncertainty, given divergent incentives and interdependence of the players, how can we structure the game to provide optimal social gains? Game theory models the way legal rules can influence strategic actors by altering the information structure of the game, players' strategies, or their payoffs.[376] There is no duty to disclose everything a director knows or learns, or every business risk, but there is a duty to truthfully report (at least periodically) on the financial status of the firm. Monitoring whether the firm is doing so, however, is difficult and leads to the possibility of opportunism.[377] Misleading financial disclosures are the leading source of shareholder claims, and they appear to be increasing at the same time as claims related to mergers, acquisitions and divestitures have been more than halved.[378]

---

discretion in the number of charges brought has a greater actual effect on penalties. *See* Perino, *supra* note 286, at 686 (observing that the "penalty enhancements are unlikely to deter corporate crime to any greater degree than current provisions").

372. Tsebelis, *supra* note 370, at 356 (explaining that such announcements must be considered part of the enforcement strategy).

373. *See id.* (describing the "evolutionary model which produces cycles of crime as well as cycles of law enforcement" and explaining that even if the crime rate initially goes down in response to statements about increased enforcement, decreased crime will provoke decreased monitoring, and ultimately the crime rate will rise again).

374. *Id.* at 360.

375. Ayres & Braithwaite explained that the "successful pursuit of cooperative regulation and maximum compliance with the law is predicted by: use of tit-for-tat strategy, access to a hierarchical range of sanctions and a hierarchy of interventionism in regulatory style (the enforcement pyramid); and how extreme in punitiveness is the upper limit of the range of sanctions." AYRES & BRAITHWAITE, *supra* note 100, at 65.

376. For a comprehensive discussion of game theory models, see generally BAIRD ET AL., *supra* note 6.

377. *See* MILGROM & ROBERTS, *supra* note 20, at 167 (discussing the concept of moral hazard, the "form of postcontractual opportunism that arises because actions that have efficiency consequences are not freely observable and so the person taking them may choose to pursue his or her private interests at others' expense).

378. *See* TILLINGHAST-TOWERS PERRIN, 2001 DIRECTORS AND OFFICERS LIABILITY SURVEY

Well run corporations should establish a corporate policy addressing risk, and adequately inform themselves about those risks, as well as delineate policies about when such risks must be disclosed.[379] In terms of what should be disclosed, information that is—or ought to be, in a well-run company—before the directors and officers includes financial information, current business developments, and future plans. This is the same information that should be before the investors and other stakeholders, such as employees.[380] Making sure that this information gets out to the market is equally important.[381] There is strong evidence of the link between financial statement fraud and weak corporate governance.[382] But mandating corporate compliance programs, such as the system under the Sarbanes-Oxley Act, imposes high costs without any indication of their effectiveness. A far more efficient solution would be to empower the tripartite structure of the enforcement pyramid through a system of director liability, actively enforced by the SEC and private litigation. Although increasing the penalties on risk-averse decisionmakers may impose excessively high social costs by stifling innovation,[383] this reasoning does not apply to risk-

---

7 (2002) (reporting that 38.8% of all shareholder claims involved inadequate or inaccurate financial reporting, up from 34.5% in 2000 and 19.9% in 1990. At the same time, the percentage of shareholder claims relating to merger, acquisition and divestiture activity declined from 40% in 1990 to 18% in 2001). *But see* Patricia M. Dechow & Douglas J. Skinner, *Earnings Management: Reconciling the Views of Accounting Academics, Practitioners, and Regulators*, 14 ACCOUNTING HORIZONS 235, 244 (2000) (expressing uncertainty as to whether the kind of financial reporting that crosses over into fraud is increasing or just increasingly visible).

379. *See, e.g.*, Dennis R. Dumas, *Targeting the Board*, BUS. L. TODAY, June 6, 1997, at 30 (discussing the importance of board policies regarding risk disclosure).

380. *See* COX ET AL., *supra* note 175, at 17 (explaining that the indirect costs of mandatory disclosure, such as liability and erosion of competitive advantage are minuscule compared to the uncertainty and delay of compliance with integrated disclosure for new offerings, and suggesting that the solution is a company registration process coupled with the elimination of Section 11 liability).

381. Brokers' duties under the United States federal securities laws include an affirmative obligation to have a reasonable basis for any recommendation. Mandatory disclosure gives brokers such a basis. Recognizing the importance of brokers' recommendations, they may be liable to the retail investor for any discrepancy in the information the issuer has released to the public and its recommendation. Broker liability under § 12(2) of the 1934 Act is premised on any misstatements in the broker's recommendation, together with the imposition of a suitability requirement.

382. *See* Patricia M. Dechow et al., *Causes and Consequences of Earnings Manipulations: An Analysis of Firms Subject to Enforcement Actions by the SEC*, 13 CONTEMP. ACCT. RES. 1 (1996) (documenting links between financial statement fraud and weaker governance structures such as an insider-dominated board, a CEO who is the company founder, CEO who is chairman of the board, and the absence of an audit committee); *see also* Mark S. Beasley, *An Empirical Analysis of the Relation Between the Board of Director Composition and Financial Statement Fraud*, 71 ACCOUNT. REV. 443 (1996) (similar conclusions about the presence of insiders, but finding that an audit committee did not affect the probability of financial fraud).

383. Richard A. Posner, *Rational Choice, Behavioral Economics, and the Law*, 50 STAN. L.

prone managers of corporations on the brink of insolvency.[384]

To prevent director and regulatory capture, third parties who are public-regarding must be able to enforce compliance with regulation.[385] The costs of capture are increased by having third party enforcement. In addition, this tripartite enforcement structure pushes the evolution of cooperation by increasing the randomness of enforcement, similar to the firm-but-fair strategy. Recall that firm-but-fair is a strategy similar to tit-for-tat, except that it has the ability to prevail in an environment of defectors by starting with cooperate as a strategy, defecting in the next round if the other player defected in the first round, is "more wary of resuming cooperation after a round of mutual defection, and does so only with a certain probability, which depends on the precise payoff values and the expected interaction length."[386] It is more random, less predictable, and a good strategy to prevent exploitation.

Investors are the logical choice for this third party enforcement. Empowering investors avoids the problem that Ayres and Braithwaite recognized in their tripartism structure in which the public interest groups that they recommend as the third players in the regulation game are public-regarding, but not firm profitability-regarding.[387] Investors care about profitability as well as transparency.

If directors face liability, they have a personal stake in becoming informed monitors of corporate financial developments.[388] Regulatory reluctance to pursue directors for securities fraud,[389] may stem from the fear that the specter of liability will scare able directors away from service.[390] In order to keep the ranks of public firms' directors from being decimated after *Smith v. Van Gorkom*,[391] for

---

REV. 1551, 1562-63 (1998).

384. *See* Macey, *supra* note 313, at 338 (suggesting that because "increasing the probability of detecting criminals is costly, the optimal deterrence scheme may involve keeping the probability of detection low and the penalties high").

385. AYRES & BRAITHWAITE, *supra* note 100, at 71.

386. Nowak et al., *supra* note 95, at 18.

387. AYRES & BRAITHWAITE, *supra* note 100, at 71.

388. TILLINGHAST-TOWERS PERRIN, *supra* note 378, at 4, 7. D&O policies usually have three separate components, corporate reimbursement coverage, which covers the organization's indemnification responsibilities, entity coverage for claims against the organization, and personal coverage for directors and officers for situations that are not covered by the indemnification statutes. Although the average corporate reimbursement flat deductible was $418,000 for U.S. insured for-profit corporations in 2001, 96% of the U.S. survey participants had no personal coverage deductibles up from 50% of those surveyed in 1990. *Id.* at 5.

389. GAO Report, *supra* note 1, at 20 (noting how few cases are brought against directors).

390. Many articles were written after the *Van Gorkom* decision about the supposed flight of directors from service. The evidence appears to be entirely anecdotal, however. *See, e.g.*, Roberta Romano, *Corporate Governance in the Aftermath of the Insurance Crisis*, 39 EMORY L. J. 1155, 1156 (1990) [hereinafter Romano, *Corporate Governance*] (discussing periodical articles about directors leaving in droves).

391. 488 A.2d 858 (Del. 1985).

example, many states enacted indemnification statutes.[392] These statutes typically permit shareholder approved charter amendments that either eliminate or limit directors' liability for negligence.[393] Thus, directors' exposure for negligence is minimal.[394] Reckless or wrongful conduct that may expose directors to significant liability,[395] however, and that is precisely the kind of behavior at stake in the Enron, WorldCom and other recent corporate financial fraud debacles. The problem is two-fold: lack of political will to enforce regulation, and private litigation barriers. The reputational costs associated with losing lawsuits are an important deterrent, even if the costs are paid by an insurer.[396]

Insurance is a way of reducing the costs of risk bearing when people are facing statistically independent risks.[397] Relying on insurance may be problematic. There is the well known problem of moral hazard; the directors may undertake more risk if insurance will bail them out.[398] In addition, insurance

---

392. *See, e.g.*, Roberta Romano, *What Went Wrong with Directors' and Officers' Liability Insurance?*, 14 DEL. J. CORP. L. 1, 24, 29 (1989) [hereinafter Romano, *What Went Wrong*] (noting the "strong, critical reaction to the [*Smith v. van Gorkom*] decision by boards, commentators and the Delaware legislature" resulting in a majority of states enacting indemnification statutes although "the decision did not alter any substantive liability rule"). The indemnification statutes may have proved popular with shareholders because they were perceived as "eliminating a class of lawsuits where insurance payouts defray legal costs rather than compensate shareholders, and any deterrent effect is quite problematic." *See also* Romano, *Corporate Governance*, *supra* note 390, at 1156 (concluding that shareholder derivative suits were the impetus for the widespread adoption of charters indemnifying directors for negligence).

393. Romano, *What Went Wrong*, *supra* note 392, at 34. Insurance may be preferable to indemnification even if they cover roughly the same exposure because if the corporation becomes bankrupt, it will be unable to pay litigation claims, while the insurance will be unaffected.

394. *See* Romano, *Corporate Governance*, *supra* note 390, at 1160-61 (observing that "over 90% of a random sample of 180 Delaware firms adopted a limited liability provision within one year of the statute's enactment").

395. *See id.* at 1161 ("[C]lass actions alleging federal securities law violations tend to generate larger recoveries than derivative suits.").

396. *See, e.g.*, Noel O'Sullivan, *Insuring the Agents: The Role of Directors' and Officers' Insurance in Corporate Governance*, 64 J. RISK & INS. 545, 546 (1997) (explaining arguments in favor of insurance).

397. *See* MILGROM & ROBERTS, *supra* note 20, at 211 (explaining that "sharing independent risks reduces the aggregate cost of bearing them"). Independent risks are those that are unrelated to each other, for example the size of the state lottery and the level of the Dow Jones Industrial average.

398. *See id.* at 174-76 (discussing the perverse effects of the moral hazard problem posed by the juxtaposition of federal insurance with rate competition in the context of the savings and loan disaster of the 1980's). The savings and loan crisis of the 1980s involved moral hazard not only with respect to the S&L owners who gambled with their depositors' money (knowing that insurance would cover their losses), but also the depositors (who relied on insurance rather than monitoring the banks), and the legislature (which raised the amount of available insurance, attracting large

is subject to cyclical availability that is not well understood.[399] Exclusions may limit coverage in unanticipated ways.[400] The premiums for Directors and Officers Liability Insurance are increasingly expensive.[401] These problems are not insurmountable, however.

Increased monitoring and verification are one solution to the moral hazard problem.[402] That was the role that auditors, the stock exchanges (sro's), rating agencies and government regulators were supposed to play. While each of these play a role in developing competing sources of information, they each have their own interests that diverge from the goal of investors in monitoring the directors. Insurance, on the other hand, has a stake in the monitoring process that is better aligned with that of the shareholders.[403]

The "tall enforcement pyramid" of Ayres and Braithwaite, using escalating regulatory sanctions in a tit-for-tat strategic game,[404] solved the problem of discovering information in large bureaucracies by having internal inspectors more familiar with the workings of the corporation than outsiders could be.[405]

Directors' and officers' liability insurance ("D&O" insurance) typically covers the costs of lawsuits against directors and officers brought by shareholders and third parties, as long as there is neither an admission nor a judicial finding of bad faith.[406] Insurance costs and premiums will reflect litigation and business risks, and insurers demand information from firms to assess these risks. Claims and notifications of suits give the insurer an opportunity to examine the aspects

---

deposits to the S&L's without increased regulatory monitoring). *Id.* at 176.

399. *See, e.g.*, Tillinghast-Towers Perrin, 2001 Directors and Officers Liability Survey Summary at 3 (stating that "the firming of the D&O market—even the sharp increase seen by some sectors—does not signal a return to crisis conditions similar to those of the mid-1980's" and noting that there is "less dependence on a small group of reinsurers now than 16 years ago"). *See id.* at 9.

400. For example, although there is no "standard" directors and officers liability insurance policy, most policies have exclusions for self-dealing, and the number of potential policy exclusions has increased since 1984. *Id.* In addition, court interpretations of policy exclusions may create some uncertainty.

401. *See id.* at 2 (noting that "[n]early all segments in the U.S. saw sharp increases in premiums as well as more stringent underwriting").

402. *See* MILGROM & ROBERTS, *supra* note 20, at 186 (one remedy for moral hazard problems is increased monitoring and verification).

403. *See, e.g.*, John E. Core, *The Director's and Officer's Insurance Premium: An Outside Assessment of the Quality of Corporate Governance*, 16 J. L. ECON. & ORG. 449, 449 (2000). (finding "a significant association between D&O premiums and variables that proxy for the quality of firms' governance structures").

404. *See* AYRES & BRAITHWAITE, *supra* note 100, at 38-39.

405. *See id.* at 105 (citing Braithwaite's studies of the pharmaceutical industry in which managerial inside knowledge of people and processes permitted him effective quality controls).

406. *See* Core, *supra* note 403, at 450 (describing D&O liability coverage). Corporate coverage reimburses the firm when it indemnifies its officers and directors, and personal coverage provides officers and directors with direct coverage if the corporation does not. *Id.* at 453-54.

of corporate governance giving rise to the dispute, and thus provide for external monitoring of the firm.[407]

Empowering a third group that has a direct stake in the interactions is a way out of the capture conundrum in which directors' and regulators' sucker payoffs are less than their punishment payoffs. The problems of cyclical availability and what firms may consider to be exorbitant pricing can be solved by self-insurance. A self-regulatory group for large publicly held corporations specifically focused on detecting financial fraud may solve the moral hazard problem of diffuse responsibility for financial reporting, as well as cyclical availability and escalating costs that are unrelated to risk.

The structure I propose would be a self- insurance group that consists of financial specialists. They would do both regular and spot inspections, and advise the firm of any problems discovered in the audit. The information they uncovered would be confidential.[408] Because the insurance rates would depend on the compliance of the firms, the firms have a stake in complying with the self-regulators. And the insurance group has its own interests—keeping liability down—motivating it to do a thorough inspection. Because the majority of the recent corporate debacles appear to have involved financing vehicles and capitalization decisions, the insurance would be limited to financial disclosures.

CONCLUSION

Evolutionary game theory and empirical studies of cognition not only challenge some of the fundamental assumptions of law and economics, they also provide insights into the role of law in shaping optimal social interactions. Socially efficient norms will not necessarily prevail without assistance. Reciprocity is a key to human interactions, but evolutionary game theory demonstrates the importance of structuring initial conditions and providing coordinating signals to achieve socially optimal payoffs. Because not all circumstances permit socially efficient norms to prevail, relying on market forces to channel behavior is evolutionarily shortsighted. Regulation and enforcement are important components of well functioning capital markets. Enforcement efforts (private and public) have been dramatically curtailed, however. Investors' legal protections have been shrinking. Legislative reform and judicial activism have both eaten away at core investor protection principles. In the wake of this trend, a series of spectacular corporate debacles has made headlines around the world. From an international perspective, accommodating regulatory needs while encouraging harmonization of a global marketplace demands sensible minimal regulation coupled with shareholders empowered to police fraudulent statements and omissions.

---

407. *See* O'Sullivan, *supra* note 396, at 545 (concluding that in large publicly held companies, D&O insurance performs a monitoring function).

408. The SEC does not require firms to disclose anything about their D&O insurance to their shareholders. Core, *supra* note 403, at 475. Thus, any information uncovered by the self-investigatory body should remain similarly confidential.

Large public corporations, with their bureaucratic diffusion of responsibility, pose an immense challenge to efficient markets that depend on a free flow of accurate information for their well being. The Enron implosion is an illustration of the problems that such diffusion of responsibility can create. The immediate congressional reaction to Enron, however, enacting the Sarbanes-Oxley Act, was neither necessary nor sufficient to solve these problems. The corporate governance provisions of the new legislation, although far from novel, are misconceived. The corporate governance provisions do little to change existing law, while imposing high costs on corporate shareholders. Moreover, although the concept of a corporate/ government partnership in fighting corporate crime has gained academic and political currency, there are good reasons to doubt the efficacy of corporate compliance programs as a partner in crime control.

The idea that companies must conduct their business with as much openness as possible is consonant with ideals of corporate democracy[409] and with the assumption that people make better decisions if they have more information. Increased knowledge decreases uncertainty.[410] The economic meaning of information is not only data, but also the web of social practices through which the data has meaning.[411] At a very minimum, government's task is to ensure that there is an appropriate macroeconomic climate for decisionmaking.[412]

Directors should not be able to evade liability for the abdication of their oversight duties. Functional monitoring is a prerequisite for a sound economy. Voluntary acceptance of rules that promote participants' objectives is undoubtedly preferable to sanctions as an economic solution to achieving cooperative behavior.[413] It is certainly cheaper. But evolutionary game theory posits that stabilizing cooperative interactions requires would-be defectors to face the threat of sanctions and "that those who are charged with identifying defectors

---

409. *See* Stephen Labston, *Bush Doctrine, Lock Em' Up*, N.Y. TIMES, June 16, 2002, § 3, at 12 (quoting Donald C. Langevoort) ("[T]he broader view is that the investor needs not only a sense of protection from bad apples [individual miscreants], but that companies must conduct themselves with an eye toward more openness.").

410. *See* MARIO BUNGE, FINDING PHILOSOPHY IN SOCIAL SCIENCE 83 (1996) (arguing that the larger the number of variables in a particular problem, and the less is known about the variables' interrelationships, the more complex the situation becomes, and the less relevant prior knowledge becomes).

411. *See* Gerhard Roseger, *Aspects of Uncertainty and Complexity in Technologies and Technosystems*, EVOLUTION AND PROGRESS IN DEMOCRACIES 123, 140 (Johann Gotschl ed., 2001) (arguing that innovation requires conditions for the diffusion of existing knowledge, and that in the United States, the willingness to "suspend belief in competitive markets as the primary source of all desirable innovations" resulted in creative technological innovations, but that these innovations nonetheless were motivated by market signals).

412. *See id.* at 125 (discussing the importance of government in shaping conditions that stimulate cooperation).

413. McClennon, *supra* note 149, at 183 (arguing that a commitment to rules is instrumentally rational as a way of solving coordination problems).

and carrying out such sanctions be sufficiently motivated to do so."[414] In repeated interactions, informal norms of reciprocity may emerge, but only if participants expect that defection will be met with retaliation at the next iteration of the game.[415] Increased reporting and punishment of defectors yield increased cooperation if others in the community, who are not necessarily co-players, also retaliate.[416] Thus, although insurance, like compliance programs, is a cost that will be borne by the shareholders, it is more likely to be effective in deterring corporate misconduct. Because the Sarbanes-Oxley Act fails to recognize or accommodate the interactive strategic processes of director decisionmaking, it is not likely solve the problems it set out to address, and will have little effect on deterring or preventing corporate misconduct. In sum, insuring that directors exercise their oversight functions is vital for a healthy economy.

414. *Id.* at 209-10 n.55.

415. *Id.* at 200.

416. *Id.*

Indiana Law Review

Volume 37 **2003** **Number 1**

NOTES

REEMPLOYMENT RIGHTS UNDER THE UNIFORM SERVICES EMPLOYMENT AND REEMPLOYMENT ACT (USERRA): WHO'S BEARING THE COST?

ANTHONY H. GREEN*

INTRODUCTION

Imagine phone calls coming to lawyers, doctors, mechanics, assembly-line workers or supervisors, production managers, dish washers, or chefs, the list goes on. Each phone call tells the individual that he or she must pack her bags, kiss her family good-bye, because there is yet another military mission to an undisclosed location for an indefinite amount of time. In some instances the phone call, instead of ordering the individual involuntarily away, gives the person a choice between volunteering his services for thirty to forty-five days several times a year for the next few years or being involuntarily ordered away for an indefinite period of time. This is the sacrifice many men and women choose to make on a daily basis as Guard personnel or Reservists.

The employer of the Guard person and Reservist, however, has no choice in the matter but is statutorily and judicially forced to maintain their positions of employment for indefinite periods and bear the cost without any compensation. This Note will address the fact that the cost of maintaining a large professional military has been externalized and shifted from the government to the employer. This shift in cost calls for either a legislative or judicial remedy to compensate for the unfair burden placed on the employer because of the language and interpretation of the USERRA.

In a complex and highly dependent world in which information is immediate, the U.S. military is continuously called to end, resolve, and head off potential conflicts. As the size of the U.S. military decreased after the Cold War, the military's activity increased. As Air Force Lieutenant General Thomas McInerney explained, "Our deployments have gone up three to four hundred percent, and we have a force that is 40 percent smaller than we had in 1988-89."[1] Therefore, the gap between supply and demand of the U.S. military has been filled to a large extent by the Guard and Reservists—the citizen soldiers.

---

* J.D. Candidate, 2004, Indiana University School of Law—Indianapolis; B.S., 1989, United States Air Force Academy, Colorado Springs, Colorado.

1. Brian Cabell, *U.S. Faces Challenge Recruiting Reservists* (Sept. 12, 2000), *at* http://www.cnn.com/2000/US/09/11/us.reservists/index.html (last visited July 7, 2003).

In 1994, Congress enacted the Uniformed Services Employment and Reemployment Rights Act "to expand, codify, and clarify the employment rights and benefits available to veterans and employees."[2] In response to the end of the Cold War and the restructuring of the U.S. military, Congress created a statute to encourage non-career military service by minimizing the disadvantages to civilian careers and employment by prohibiting discrimination against individuals because of their military service.[3] Congress borrowed several concepts from other federal employment discrimination statutes to provide for prompt reemployment of the service member upon completion of his or her service.[4]

The USERRA reflected a shift in the nation's defense policy with a new reliance on the citizen soldier. This new force structure was deemed the "Total Force Policy," which saw the Reserve and Guard components as an integral part of the military resources on hand at any given time.[5] "The Total Force Policy called for an increased reliance on the reserves and was implemented in an effort to make training 'more meaningful' for these components and boost military manpower."[6] Currently, the 1.3 million men and women who serve in the Reserve and Guard make up nearly half of the U.S. Armed Forces.[7]

Initially, the shift in defense policy appeared to be a solution to having a well-trained force on call without the cost of maintaining a large professional standing military. However, with mass activations in 1998 for *Operation Allied Force for Kosovo*, and *Operations Enduring Freedom* and *Noble Eagle* following September 11, 2001, the citizen soldier has been called to duty for extended periods several times in the last few years. "This is the first time since the Vietnam War and creation of the all-volunteer military in 1973 that reserve troops have been asked to stay on active duty" for such periods.[8] Furthermore, other Reservists and Guardsmen, to avoid being called up involuntarily for an indefinite amount of time, volunteer for thirty to forty-five day rotations several times a year. Major General Czekanski described the plight of the Guardsmen and Reservist by rejecting the model of the Guard and Reserves as a one weekend

---

2. Eve I. Klein & Maria Cilenti, *When Duty Calls: What Obligations Do Employers Have to Employees Who Are Called to Military Service?*, 73-DEC N.Y. ST. B.J. 10, 10 (2001); *see also* Lieutenant Colonel H. Craig Manson, *The Uniformed Services Employment and Reemployment Rights Act of 1994*, 47 A.F. L. REV. 55, 59 (1999).

3. *See* Klein & Cilenti, *supra* note 2, at 10.

4. Manson, *supra* note 2, at 59.

5. Andy P. Fernandez, *The Need for the Expansion of Military Reservists' Rights in Furtherance of the Total Force Policy: A Comparison of the USERRA and ADA*, 14 ST. THOMAS L. REV. 859, 861 (2002).

6. *Id.*; *see also* Samuel W. Asbury, *A Survey and Comparative Analysis of State Statutes Entitling Public Employees to Paid Military Leave*, 30 GONZ. L. REV. 67, 71 (1994).

7. U.S. Department of State, *Bush Orders 50,000 U.S. Reservists to Active Duty* (Sept. 14, 2001), *at* http://usinfo.state.gov/topical/pol/terror/01091403.htm (last visited Sept. 14, 2001).

8. Reuters, *U.S. Reserves May Stay on Duty for Second Year* (Aug. 26, 2002), *at* http://www.ledger-enquirer.com/mld/ledgerenquirer/3942819.htm (last visited July 9, 2003).

a month, two weeks a year program.[9] The General explained that Reservists need to expect to give at least sixty days a year above and beyond this commitment, and if the volunteerism is not there, rolling activations of up to 120 days may be in order.[10]

The USERRA was created to encourage volunteerism that would allow the country to maintain a large force of citizen soldiers that could be trained and ultimately sent to war.[11] Upon the soldiers' return from war, they would come home to the job that they had left. However, the current force structure and subsequent military policy is dependent on the Reservists and Guardsmen not only in times of war, but also in continuous military operations occurring around the world. Ohio congressman David Hobson explained that people who joined the Reserves and Guard in the past did so with an understanding that periods of being called up to duty would be "relatively short."[12] The congressman described the current situation, in which "[p]art-time reservists are being turned into full-time soldiers and airmen through extended and unpredictable active-duty assignments."[13] This dependence on Reservists and Guardsmen in the Total Force Policy with continuous world-wide military operations has created a situation where employers are faced with a revolving door, unsure when, where, and how often their employees will be voluntarily or involuntarily taken from them.

Courts have found, under other employment laws, such as the Americans with Disabilities Act (ADA) and the Family and Medical Leave Act (FMLA), that revolving door situations where employees attempt to take periods of indefinite leave create an unreasonable burden for employers.[14] This is similar to the situation employers find themselves in under the USERRA in today's world. If the citizen soldier is called to go to war to defend his or her country and way of life, employers reap the benefits just the same as every other citizen. However, with today's U.S. force structure and consequent political and military policy, the citizen soldier is being called for continuous military operations with the cost disproportionately placed on the employer.

The savings in relying heavily on a part-time citizen soldier as opposed to a full-time professional soldier are tremendous. The Guard and Reserves "allow the nation to nearly double its armed forces . . . while accounting for just 8.3% of the defense budget."[15] The government cost savings from using Reservists instead of the professional soldier is analogous to businesses relying on

---

9. Major General James P. Czekanski, Remarks at Grissom Air Reserve Base (May 5, 2002).

10. *Id.*

11. Klein & Cilenti, *supra* note 2, at 10.

12. Greg J. Zoroya, *Citizen Soldiers Report Long Tours, Little Support*, U.S.A. TODAY, Jan. 16, 2003, at A1 (quoting Congressman David Hobson).

13. *Id.*

14. *See* Nowak v. St. Rita High Sch., 142 F.3d 999, 1004 (7th Cir. 1998); Monette v. Elec. Data Sys. Corp., 90 F.3d 1173, 1187 (6th Cir. 1996); Rogers v. Int'l Marine Terminals, 87 F.3d 755, 760 (5th Cir. 1996); Tyndall v. Nat'l Educ. Ctrs. Inc., 31 F.3d 209, 213 (4th Cir. 1994).

15. Zoroya, *supra* note 12, at A1.

temporary employees as opposed to salaried employees. However, the burden is not distributed equally but instead placed primarily on the employer of the citizen soldier. The employer is expected not only to continue its operations not knowing when and for how long their employees will be taken from them but also to keep these positions for them when they return. As a matter of policy, the government must have the option of taking its citizens to defend the country and our way of life, but the burden must be fairly distributed. This should require the government to give some type of compensation to the affected employers.

Congress explicitly declared that the purpose of the USERRA "is to encourage non career service . . . by eliminating or minimizing the disadvantages to civilian careers and employment."[16] Congress sought to prohibit discrimination against individuals because of their service in the military.[17] To help accomplish its objectives, Congress differentiated the statute into three major elements: (1) a prohibition on employment discrimination against service members; (2) reemployment rights for persons absent from employment because of military service; and (3) preservation of benefits for persons absent from employment because of military service.[18]

The second element of the USERRA—the reemployment rights, will be the focus of this Note. The legislative history of the USERRA as well as the major provisions of the USERRA that apply to reemployment rights will be addressed in Part I. In Part II, the reemployment provisions and their applicability in current case law will be discussed. Therein, the Note will explain the process of starting a claim and the administrative channels through which a claim might move. The section will then describe the burden of proof in USERRA claims by looking at cases with a showing of discrimination and cases without a showing of discrimination in which the argument for mandatory reemployment is used. Part III of the Note will begin by explaining the difference between an unreasonable accommodation and undue hardship in the ADA. Part III will next discuss two different types of cases under the ADA: (1) cases where the main issue was whether an employee could meet the burden of proof that indefinite leave was a reasonable accommodation; and (2) cases where the employee met this burden, thereby shifting the burden to the employer to prove that providing indefinite leave as an accommodation was undue hardship. Through the discussion of these two types of cases and the cases addressed in Part III, USERRA claims will be compared and contrasted with the ADA cases.

Finally, this Note will offer solutions to the mandatory requirement of reemployment, which many employers consider unreasonable, an undue hardship or both. The solutions will first suggest that legislation is the most effective but least likely solution. The next solution will contain suggestions from the analyzed case law in Part III of the Note suggesting an application of the ADA burden of proof in USERRA cases. The solutions will address that, as a matter of fairness, that the cost of maintaining a large professional force has been

---

16. 38 U.S.C. § 4301(a)(1) (2003).

17. *Id.* § 4301 (a)(3).

18. Manson, *supra* note 2, at 59 (citing 32 C.F.R. § 104.1 (1998)).

largely externalized and shifted from the government to the employer. Therefore, a remedy of compensation from the government to the employer is necessary. Finally, the solutions will show that a more practical approach may be to draw a reasonableness standard from the FMLA to use as a guideline to determine when and what type of compensation should be required.

I. USERRA

A. History

The first legislation enacted for the benefit of the deployed serviceman occurred during the Civil War when Congress passed legislation suspending any criminal or civil action against federal soldiers while serving in the Union army.[19] Later, during World War I, the army drafted a soldiers' and sailors' relief bill, which became known as the Soldiers' and Sailors' Civil Relief Act of 1918 (SSCRA).[20] The SSCRA ended when a provision in the statute was activated approximately six months after the war had ended.[21] When the country found itself on the brink of World War II over twenty years later, the SSCRA was reenacted to reassure servicemen that the lives they were leaving behind to fight the war would be, to some extent, kept in order.[22] Along with SSCRA, Congress created the Selective Training and Service Act of 1940 (STSA), which first established veterans' reemployment rights.[23] Congress anticipated the need to train a large number of civilians into the small professional military.[24] If no war occurred, these civilians would go back to their civilian occupations after training.[25] If war did break out, they would return to their jobs at the end of their service.[26] During hearings on the proposed legislation, Senator Elbert D. Thomas of Utah expressed the idea "underlying Congress's decision to grant reemployment rights to veterans."[27] Senator Thomas explained that

> [i]f it is constitutional to require a man to serve in the Armed Forces, [sic] it is not unreasonable to require the employers of such men to rehire

---

19. Amy J. McDonough et al., *Crisis of the Soldiers' and Sailors' Civil Relief Act: A Call for the Ghost of Major (Professor) John Wigmore*, 43 MERCER L. REV. 667, 669 (1992).

20. *Id.* at 670.

21. *Id.*

22. Fernandez, *supra* note 5, at 869.

23. *Id.*

24. Manson, *supra* note 2, at 56 (citing Selective Training and Service Act of 1940, Pub. L. No. 783, 54 Stat. 885, formerly codified at 50 U.S.C. § 308, repealed by Selective Service Act of 1948, Pub. L. No. 759, 62 Stat. 625).

25. *Id.*

26. *Id.*

27. Judith Bernstein Gaeta, Note, Kolkhorst v. Tilghman: *An Employee's Right to Military Leave Under the Veterans' Reemployment Rights Act*, 41 CATH. U. L. REV. 259, 264 (1991) (citing Office of Veterans' Reemployment Rights, U.S. Dep't of Labor, The Veterans' Reemployment Rights Handbook, at 1-2 (1988)).

> them upon the completion of their service, since the lives and property of the employers as well as everyone else in the United States are defended by such service.[28]

The U.S. Supreme Court agreed with Senator Thomas' belief "when the Court held that an employer should not penalize a veteran who served in the Armed Forces when the veteran returned to his civilian job after an absence for military duty."[29]

In 1948, the STSA was reenacted as the Selective Service Act (SSA) without meaningful changes to veterans' reemployment rights.[30] Shortly thereafter, the SSA was modified and renamed the Universal Military Training and Service Act (UMT).[31] The purpose of the legislation "was to support the conscription-based force management policies that existed"[32] after the Cold War. "In 1967, the UMT was renamed the Military Selective Service Act of 1967 (MSSA)" without the UMT reemployment provisions being changed.[33] "The MSSA was amended in 1968 to give protection to reservists and National Guardsmen against discrimination after their reemployment because of their military" duty.[34] Finally, in 1974, Congress repealed the MSSA by enacting the Vietnam Era Veterans' Readjustment Assistance Act and recodifying the MSSA's provisions.[35] At the end of the Vietnam War, large numbers of service members separated from the military as involvement in Vietnam came to an end.[36] "Additionally, the draft had ended and the nation was transitioning to a peacetime all-volunteer force."[37] "Employment protection was important in luring the potential one-term volunteer . . . and to induce separating members to continue to serve in the [Guard and] reserve forces."[38]

Congress amended the reemployment legislation several times and finally ended up with the present reemployment rights legislation, Uniformed Services Employment and Re-employment Rights Act of 1994 (USERRA).[39] The USERRA, along with the end of the Cold War, which caused another draw down

---

28. *Id.* at 264 n.26 (quoting remarks of Sen. Thomas, 123 CONG. REC. 10, 573 (1940)).

29. *Id.* at 264 (citing Fishgold v. Sullivan Drydock & Repair Corp., 328 U.S. 275, 284 (1946)).

30. *Id.* (citing Selective Service Act of 1948, Pub. L. No. 759, 62 Stat. 604).

31. *Id.* (citing 1951 Amendments to the Universal Military Training and Service Act, Pub. L. No. 51, §§ 1(a), 3(w), 65 Stat. 75, 86-87 (1951) (formerly codified at 50 U.S.C. § 459 (1948)).

32. Manson, *supra* note 2, at 56.

33. Gaeta, *supra* note 27, at 265 (citing Military Selective Service Act of 1967, Pub. L. No. 90-40, § 1(a), 81 Stat. 100).

34. *Id.* (citing Armed Forces, Reserve Components, Pub. L. No. 90-491, § 1(B), 82 Stat. 790 (1968)).

35. *Id.* (citing 38 U.S.C. §§ 2011-2026 (1976)).

36. Manson, *supra* note 2, at 57.

37. *Id.*

38. *Id.* (citing S. Rep. No. 93-907, at 110 (1974)).

39. *Id.* at 57-58.

of active duty forces, combined to bring about another shift in national defense policy.[40] Furthermore, in the post-Cold War era, the United States faces different security threats and geographic positioning as fewer military personnel are stationed at overseas military bases than before.[41] The result of these facts is that both Reserve and Guard personnel are called on to deploy for varying and often indefinite periods of time.[42]

B. Reemployment Rights Provisions

Under the USERRA, a person covered includes one "who is a member of, applies to be a member of, performs, has performed, applies to perform, or has an obligation to perform service,"[43] "on a voluntary or involuntary basis in a uniformed service under competent authority."[44] Virtually all areas of military service are covered, including "active duty, active duty for training, initial active duty for training, inactive duty training, full-time Natural Guard duty," and absences for fitness examinations.[45] USERRA covers all employers in the United States regardless of how many employees they have and whether they are a person, institution, organization, or other entity to whom the employer has delegated the performance of employment-related responsibilities.[46]

USERRA is basically separated into three sections, the first two of which this paper will address. The first prong of the USERRA "is directed at prohibiting discrimination against . . . service members."[47]

> The presence of discrimination against the service member is detected when his military membership, service, or application for service is deemed a "motivating factor" in an employer's negative action, absent a showing by the employer that the actions would have taken place against that employee regardless of any such connection to military service.[48]

The USERRA states the reemployment rights in its second prong.[49] This section defines reemployment rights and requirements, defense for employers and the situations when they may refuse reemployment, and the returning employee's reemployment requirements adjusted for his or her length of duty.[50]

An employee who is called for military duty is entitled to re-employment

---

40. *Id.* at 58.
41. *Id.*
42. *Id.*
43. 38 U.S.C. § 4311(a) (2003).
44. *Id.* § 4303(13).
45. *Id.*
46. Manson, *supra* note 2, at 60 (citing 38 U.S.C. § 4303(4)(A)(1) (1998)).
47. Fernandez, *supra* note 5, at 870 (citing Manson, *supra* note 2, at 59).
48. *Id.* (citing 38 U.S.C. § 4311 (2001)).
49. 38 U.S.C. § 4312 (2003).
50. *See* Fernandez, *supra* note 5, at 870 (citing 38 U.S.C. § 4312 (1998)).

> if (1) the employee has given advance written or verbal notice to the employer, unless circumstances make it unreasonable to do so; (2) the cumulative length of absence from employment with that employer, including prior service absences, does not exceed five years, with certain notable exceptions, including service in a time of war or national emergency; and (3) the employee promptly reports to work or submits an application for re-employment within 14 or 90 days after completion of service, depending on the length of service.[51]

"An employee must also submit documentation"[52] if requested by the employer "to establish that the application is timely, that the service limitations have not been exceeded, and that the employee's separation from military service was not dishonorable."[53]

If the employee meets the above conditions and is entitled to reemployment, federal regulations require "that the returning service member is entitled to the position of civilian employment that he or she would have attained had he or she remained continuously employed by that civilian employer."[54] This provision in the USERRA is known as the "escalator" principle.[55] In the first case concerning veterans' reemployment rights decided by the Supreme Court after World War II, *Fishgold v. Sullivan Drydock & Repair Corp.*, the Court observed that the individual "does not step back on the seniority escalator at the point where he stepped off. He steps back on at the precise point he would have occupied had he kept his position continuously during the war."[56] Today, the escalator principle is subject to a reasonableness standard, where the benefit sought by the returning service member is protected if it would have accrued with reasonable certainty had the employee continued to be employed, and if the benefit was not subject to a significant contingency.[57] "Another Supreme Court decision used a [similar] 'but for' analysis, holding that the returning employee is guaranteed a status, which he was reasonably certain to acquire, but for his absence due to military service."[58]

There are limitations on the requirements of reemployment, however. An employer can avoid reemploying a citizen soldier if:

> (1) the employer's circumstances have so changed that they make re-employment unreasonable; (2) re-employment would impose undue

---

51. Klein & Cilenti, *supra* note 2, at 11 (citing 38 U.S.C. §§ 4312(a)(2), 4312(e)(C), (D) (1998); Sykes v. Columbus & Greenville Ry., 117 F.3d 287, 296-97 (5th Cir. 1997)).

52. *Id.* (citing 38 U.S.C. § 4312(f) (1998)).

53. *Id.*

54. Fernandez, *supra* note 5, at 873 (citing 32 C.F.R. § 104.3 (2001)).

55. Klein & Cilenti, *supra* note 2, at 12.

56. Manson, *supra* note 2, at 71 (citing Fishgold v. Sullivan Drydock & Repair Corp., 328 U.S. 275, 284-85 (1946)).

57. *Id.* at 72 (citing Alabama Power Co. v. Davis, 431 U.S. 581, 589 (1977)).

58. Fernandez, *supra* note 5, at 873 (citing Tilton v. Mo. Pac. R.R. Co., 376 U.S. 169, 181 (1964)).

> hardship on the employer; (3) the position the employee leaves is for a brief, non-recurrent period and there is no reasonable expectation that such employment will continue indefinitely or for a significant period; or (4) the employer had legally sufficient cause to terminate the employee at the time he/she left for service.[59]

The burden of proof as to these defenses is on the employer who would not have to allow the service member employee to return to that job if the employer has a successful defense.[60]

The "USERRA [also] requires employers to use reasonable efforts to reemploy those service members in that position they would have been reasonably certain to attain, but for the absence caused by such military duty."[61] If the returning service member is unqualified for the position to which he returns, the employer must use reasonable efforts to train that individual.[62] Congress added this reasonable accommodation as a provision to help in "'eliminating or minimizing the disadvantages' to civilian careers."[63] The USERRA approach of reasonable efforts is similar to the ADA provision of reasonable accommodations.[64] In either Act, "[t]hese reasonable efforts or reasonable accommodations are exhausted . . . if the employer can produce evidence that the reemployment or reasonable efforts/accommodations impose an undue hardship upon the employer."[65] "These Acts define 'undue hardship' as actions requiring significant difficulty or expense to the employer considering such factors as the nature and cost of the reasonable efforts or accommodations and the overall resources of the business, facilities, and type of operations involved."[66]

The USERRA and the ADA define undue hardship identically and provide "the same factors for the determination of its presence except that the USERRA substitutes the term 'employer' for 'covered entity.'"[67] For both statutes, "the measure of an undue hardship may be equated to the financial costs associated with the reasonable efforts or accommodations in relation to the overall financial resources of, persons employed at, effect of expenses and resources, or the impact otherwise of the action on the facility's operation."[68]

A difference between the USERRA and other employment discrimination

---

59. Klein & Cilenti, *supra* note 2, at 11 (citing 38 U.S.C. § 4303(15) (2000); 38 U.S.C. § 4312(d)(1)(C) (2000); 42 U.S.C. § 12111(10)(A), (B)(i-iv) (1994 & Supp. V 1999); Jordan v. Jones, 84 F.3d 729, 732 (5th Cir. 1996)).

60. Manson, *supra* note 2, at 70 (citing 38 U.S.C. § 4312(d)(2) (1998)).

61. Fernandez, *supra* note 5, at 873 (citing 38 U.S.C. § 4313(a)(1)(A), (a)(2)(A) (2001)).

62. *Id.* at 874 (citing 38 U.S.C. § 4313(a)(1)(B), (a)(2)(B) (2001)).

63. *Id.* (quoting 38 U.S.C. § 4301(a)(1) (2001)).

64. *Id.* (citing 42 U.S.C. § 12111(9) (2001)).

65. *Id.* (citing 38 U.S.C. § 4312(d)(1)(B) (2001); 42 U.S.C. § 12112(b)(5)(A) (2001)).

66. *Id.* (citing Manson, *supra* note 2, at 75).

67. *Id.* at 882 (citing 38 U.S.C. § 4303(4)(a) (2001); 42 U.S.C. § 12111(10)(B)(iii) (2001)).

68. *Id.* (citing 38 U.S.C. § 4303(15)(B) (2001); 42 U.S.C. § 12111(10)(B)(ii) (2001)).

statutes is that the USERRA provides a person who is reemployed after a long-term military leave with short-term protection against discharge "except for cause."[69] This provision gives "an otherwise 'at-will' employee the equivalent of contractual protection against termination."[70] This protection covers the serviceman "for (1) one year after the date of re-employment, if military service was more than 180 days, and (2) 180 days after the date of re-employment, if military service was more than 30 days but less than 181 days."[71]

II. CURRENT APPLICATION

A. The Process of Enforcing USERRA

Before seeing how the USERRA is applied, it is helpful to understand how an employee begins the process of invoking the USERRA. Citizens who feel their rights have been violated under the USERRA can "file a complaint with the Veterans' Employment Training Service (VETS) of the U.S. Department of Labor."[72] The Secretary of Labor can use a subpoena to require "the attendance and testimony of witnesses and production of documents relating to any matter under investigation."[73] If the Secretary of Labor decides that a violation did occur, the Secretary will first make an effort to reach a "voluntary resolution" with the employer so that USERRA requirements are satisfied.[74] If the VETS does not settle the complaint, the individual may present the complaint to the Attorney General for potential court action.[75] "The statute does not mandate, however, that the service member proceed through either the Secretary of Labor or the Attorney General."[76] Therefore, unless the Attorney General agrees to prosecute a case, an employee can commence a private action against the employer under the USERRA at any time.[77]

When viewing how the USERRA is applied, one must keep in perspective the first two sections of the Act. The first section focuses on the discrimination by the employer against the employee. The second section deals with the reemployment rights of the employee regardless of discrimination. Therefore, the key to the application of the USERRA is to understand the burden of proof in relation to these two different sections of the USERRA.

---

69. Klein & Cilenti, *supra* note 2, at 10.

70. *Id.* at 14.

71. *Id.* (citing 38 U.S.C. § 4316(c) (2000)).

72. Thomas E.J. Hazard, *Employers' Obligations Under the Uniformed Services Employment and Reemployment Rights Act*, 31 COLO. LAW. 55, 58 (citing 38 U.S.C. §§ 4321-4322 (2000)).

73. Klein & Cilenti, *supra* note 2, at 17 (citing 38 U.S.C. § 4326 (2000)).

74. *Id.*

75. Hazard, *supra* note 72, at 58 (citing 38 U.S.C. § 4323(a)(1) (2000)).

76. Klein & Cilenti, *supra* note 2, at 17 (citing 38 U.S.C. § 4323(a)(2) (2000)).

77. *Id.*

B. Burden of Proof with Showing of Discrimination

The USERRA was enacted, to a large extent, to overrule the U.S. Supreme Court's decision in *Monroe v. Standard Oil Co.*[78] In *Monroe*, the Court held that under the Veterans' Reemployment Rights Act of 1968, an employer violated the veterans' rights laws only where the employee could show that his Reserve status was the "sole motivation" for the adverse action taken against the employee.[79] The USERRA overruled *Monroe* by stating that a violation occurs even when a person's military service is not the sole factor but is a "motivating factor" in the discriminatory action.[80]

The USERRA applies the standard of proof established in *National Labor Relations Board v. Transportation Management Corp.*[81] The employee has the initial burden of showing by a preponderance of the evidence that the employee's military service was "a substantial or motivating factor in the adverse [employment] action" under the "but for" test.[82] The employee does not need to show that military status was the "sole" reason for the employment action, only that it was one of several "factors that a truthful employer would list if asked the reasons for its decision."[83] The employer's discriminatory motivation or intent may be found by either direct or circumstantial evidence.[84]

If an employee demonstrates by direct evidence that his military service was a substantial or motivating factor in the employer's action, then he proves his prima facie case, and the burden shifts to the employer.[85] If an employee does not have direct evidence but can demonstrate discrimination indirectly through relevant circumstantial evidence, he can still establish his prima facie case.[86] When using circumstantial evidence to establish a claim of prohibited employment discrimination, an employee

> must first establish a prima facie case by showing that: (1) he was a member of a protected group; (2) he was similarly situated to an individual who was not a member of his protected group; and (3) he was treated more harshly or disparately than the individual who was not a member of his protected group.[87]

---

78. 452 U.S. 549 (1981).

79. *Id.* at 559. *See also* 38 U.S.C. § 2021-2026 (2000).

80. Gummo v. Vill. of Depew, 75 F.3d 98, 106 (2d Cir. 1996). *See also* 38 U.S.C. § 4311(c)(1) (2000).

81. 462 U.S. 393, 403 (1983).

82. Leisek v. Brightwood Corp., 278 F.3d 895, 899 (9th Cir. 2002) (quoting *NLRB*, 462 U.S. at 401).

83. Kelley v. Me. Eye Care Assocs., 37 F. Supp. 2d 47, 54 (D. Me. 1999) (citing Robinson v. Morris Moore Chevrolet-Buick, Inc., 974 F. Supp. 571, 575 (E.D. Tex. 1997)).

84. Sheehan v. Dep't of Navy, 240 F.3d 1009, 1014 (Fed. Cir. 2001).

85. Luecht v. Dep't of Navy, No. 00-3289, 2000 U.S. App. LEXIS 27917, at *5 (Fed. Cir. Nov. 7, 2000).

86. *Id.*

87. *Id.* at *5-*6 (citing Coleman v. Dep't of Air Force, 79 F.3d 1165 (Fed. Cir. 1996)).

"If a prima facie case is shown, the burden then shifts" to the employer, who must provide evidence that supports the existence of a nondiscriminatory reason for taking the action it did.[88]

The burden of proof moves to the employer once the employee establishes his or her prima facie case. The employer must prove by a preponderance of the evidence, "that legitimate reasons, standing alone, would have induced it to take the same adverse action."[89] Thus, under the USERRA an employer may avoid liability if it can demonstrate by a preponderance of the evidence that it took the adverse action only for a valid reason, or if an invalid reason affected the adverse action, the employer must show it would have taken the same action in the absence of that invalid reason.[90] The burden then returns to the employee to show the reasons given by the employer are a ploy for discrimination.[91]

For instance, in *Hill v. Michelin North America, Inc.*, the court explained the standard of proof and allocation of burdens at work in USERRA litigation.[92] The employee alleged that Michelin disapproved of his military Reserve obligations and punished him by transferring him to a job with irregular work schedules and long workdays and then ultimately terminating his employment. The employee asserted that his supervisor looked distraught when the employee informed him of his Reserve duty. Michelin maintained that the employee was moved to the position to accommodate his Reserve duties and that the employee was fired for falsifying his timecard. Michelin demonstrated that it terminated all employees who falsified their timecards, regardless of their participation in the Reserves or military.[93] The court found that the evidence presented by the employee of the supervisor's reaction may have been enough to meet his initial burden of proof but was not enough to overcome the employer's response.[94]

Another example of burden shifting with regard to discrimination is the case of *Leisek v. Brightwood Corp.*[95] In *Leisek*, the court found that Leisek had met his prima facie burden by bringing forth evidence from which a reasonable fact finder could deduce that Leisek's Guard status was a "motivating factor" in Brightwood's decision to end his employment.[96] The USERRA allows the discriminatory motivation of the employer to be inferred from a variety of factors such as:

---

88. *Id.* at *6.

89. *Sheehan*, 240 F.3d at 1013 (citing Nat'l Labor Relations Bd. v. Transp. Mgmt. Corp., 462 U.S. 393 (1983)). *See also* Kelley v. Me. Eye Care Assocs., 37 F. Supp. 2d 47, 54 (D. Me. 1999).

90. *See* Gummo v. Vill. of Depew, 75 F.3d 98, 106 (2d Cir. 1996); *see also* Leisek v. Brightwood Corp., 278 F.3d 895, 899 (9th Cir. 2002); Robinson v. Morris Moore Chevrolet-Buick, Inc., 974 F. Supp. 571, 576 (E.D. Tex. 1997).

91. *Id.*

92. 252 F.3d 307 (4th Cir. 2001).

93. *Id.* at 314.

94. *Id.*

95. *Leisek*, 278 F.3d at 895.

96. *Id.* at 900.

> proximity in time between the employee's military activity and the adverse employment action, inconsistencies between proffered reason and other actions of the employer, an employer's expressed hostility towards members protected by the statute together with knowledge of the employee's military activity, and disparate treatment of certain employees compared to other employees with similar work records or offenses.[97]

The evidence in *Leisek* provided for an inference that Leisek's military status was a "motivating factor" in Brightwood's decision to end his employment because of the significant number of absences from work caused by Leisek's participation in the Guard.[98] "The record include[d] testimony supporting an inference that Leisek's Guard-duty absences . . . had created an increased burden for Brightwood and that it had proposed a plan that would restrict Leisek's future Guard-related absences to a period of three weeks," and that those absences would be taken away from his vacation time.[99] In addition, Leisek was told that Brightwood had ordered him to stop volunteering for ballooning events and decided not to honor any future Guard orders, except for those that it already had been given.

The burden then shifted to Brightwood, who claimed that the evidence established the affirmative defense that it would have ended Leisek's employment for his unauthorized absences, regardless of his Guard status or conduct.[100] The employer has the burden of the proof with respect to this affirmative defense.[101] There was evidence that Brightwood's corporate guidelines made unexcused absences a basis for ending employment.[102] "However, even though Leisek's unexcused absences would be a legitimate reason for terminating his employment, Brightwood ha[d] not established . . . that it would have terminated Leisek even if he had not been active in the Guard."[103]

C. Burden of Proof Without a Showing of Discrimination

In situations where an employee leaves for military duty, most courts find that the employer has a nearly absolute obligation to reemploy the returning serviceman under the second prong of the USERRA. If the returning serviceman can meet the burden of proving the elements of 38 U.S.C. § 4312, the employer must reemploy its former employee.[104] There is no burden shifting exercise such

---

97. *Id.*
98. *Id.*
99. *Id.*
100. *Id.*
101. Sheehan v. Dep't of Navy, 240 F.3d 1009, 1014 (Fed. Cir. 2001).
102. *Leisek*, 278 F.3d at 900.
103. *Id.*
104. Fernandez, *supra* note 5, at 871. *See also* 38 U.S.C. § 4312 (2000); Klein & Cilenti, *supra* note 2, at 11.

as in situations when there is a showing of discrimination.

However, in the case of *Curby v. Archon*,[105] the court explained that the proper statutory interpretation of the USERRA was to look at the specific statutory language and design of the statute as a whole.[106] The court found that § 4311 has a very broad scope that covers discrimination in initial employment, reemployment, retention in employment, and promotion, whereas § 4312 has a very narrow scope addressing only reemployment after a leave of absence for military service.[107] The court saw § 4312 as a subsection of § 4311 and "therefore conclude[d] that a person seeking relief under § 4312 must meet the discrimination requirement contained in § 4311."[108] In other words, the court found that the returning servicemen must show that the military service was a "motivating factor" in the termination of employment and then show the elements of § 4312.

Most courts have not agreed with the interpretation of the court in *Curby* and have interpreted § 4312 separately from § 4311. For instance, in *Wrigglesworth v. Brumbaugh*,[109] Wrigglesworth was forced to present his resignation to his employer while on military leave.[110] The employer refused to permit him to retain his previous level of seniority or to advance him to the level he would have reached but for his absence when he returned. The court held this to be a violation of § 4312 by reasoning that "[S]ection 4311 and Section 4312 are independent, with only Section 4311 requiring a finding of discriminatory intent."[111] The court found that the House Report and other legislative history "make[] clear that the purpose of Section 4312 was to provide an automatic right of re-employment different from the right described in Section 4311" and not dependent on proof of discrimination.[112] Furthermore, the court in *Wrigglesworth* distinguished its case from *Curby* by claiming that the facts of the two cases were different.[113] *Curby* dealt with an employee fired under the "just cause" provision under 38 U.S.C. § 4316(c) for gross misconduct. In *Wrigglesworth*, there was no application of the "just cause" provision. The court noted that the *Curby* court's interpretation of the relationship between § 4311 and § 4312 was dicta because the decision rested on the "just cause" provision.[114]

The court in *Wrigglesworth* also addressed the collective bargaining agreement between the employee and the employer waiving reemployment rights under the USERRA. The court explained that "the Supreme Court's decision in

---

105. 216 F.3d 549 (6th Cir. 2000).

106. *Id.* at 557.

107. *Id.*

108. *Id.*

109. 121 F. Supp. 2d 1126 (W.D. Mich. 2000).

110. *Id.* at 1128-29.

111. Jordan v. Air Prods. & Chems., Inc., 225 F. Supp. 2d 1206, 1208 (C.D. Cal. 2002) (quoting *Wrigglesworth*, 121 F. Supp. 2d at 1135-36).

112. *Wrigglesworth*, 121 F. Supp. 2d at 1136.

113. *Id.* at 1137.

114. *Id.*

Fishgold made clear that a [sic] employment decision denying re-employment rights cannot be cloaked in the language of a collective bargaining agreement."[115]

In another case, *Jordan v. Air Products and Chemicals, Inc.*,[116] the employer attempted the same type of defense as in *Wrigglesworth*, by trying to place the burden of proof on the employee to show that the termination of employment was a motivating factor. However, the court disagreed with that burden shifting process for slightly different reasons. Jordan had given his employer, Air Products, advance notice that he would be absent from his employment for approximately seventeen days due to his service in the Reserves.[117] Shortly after reporting back to work from his Reserve tour, Jordan was notified by Air Products that his employment was terminated, effective immediately. Jordan brought suit against Air Products for violation of the USERRA. Relying on *Curby v. Archon*, the defense claimed that the employee is required to show not only a failure to reemploy, but also that the person's military service was a motivating factor in the employer's decision under § 4312.[118] The court disagreed, finding that § 4312 creates an "unqualified right to reemployment to those who satisfy the service duration and notice requirements."[119] As the plain language of the statute makes clear, "this benefit is subject only to the defenses enumerated in § 4312, i.e., reemployment is unreasonable, impossible, or creates an undue hardship."[120]

The court explained that once the plaintiff is reemployed under § 4312, the service person is protected by §§ 4316(c) and 4311.[121] "Section 4316 provides that a person who serves for over thirty days and is reemployed under the USERRA shall not be discharged from such employment except for cause for certain time periods."[122] Under § 4311, the decision to terminate cannot be motivated by the employee's "membership application or participation in the armed services."[123]

The right to reemployment is generally considered absolute, based upon the prevailing precedent that the legislation is to be "liberally construed to protect those who have been obliged to drop their own affairs to take up the burdens of the nation."[124] However, there is a very limited exception when the reemployment is considered unreasonable or impossible. This is considered an affirmative defense for which the employer has the burden of proof.[125] For

---

115. *Id.* at 1138 (citing Fishgold v. Sullivan Drydock & Repair Corp., 328 U.S. 275 (1946)).
116. 225 F. Supp. 2d at 1206.
117. *Id.*
118. *Id.* at 1207.
119. *Id.* at 1208.
120. *Id.*
121. *Id.*
122. *Id.*
123. *Id.* at 1208-09.
124. Boone v. Lightner, 319 U.S. 561, 575 (1943).
125. Wrigglesworth v. Brumbaugh, 121 F. Supp. 2d 1126, 1136 (W.D. Mich. 2000).

instance, in *Davis v. Halifax City School System*,[126] the court held that the reemployment of an employee is considered unreasonable "where [the] reinstatement would require creation of a useless job or where there has been a reduction in the work force that reasonably would have included the veteran."[127] Nevertheless, the burden of proof required by the courts, in all practical purposes, has proven to be almost insurmountable by the employer.

III. USERRA AND ADA COMPARISON

Under the USERRA, as under the ADA, an "employer generally must make reasonable efforts to accommodate the disability or to qualify the employee for the relevant position, if service member/employee returns from military duty with a disability or is otherwise unqualified for the job the employee left."[128] The right to reemployment is almost absolute. However, like under the ADA, an employer is not required to reemploy a person or to accommodate or qualify the person if doing so would impose an undue hardship on the employer or is unreasonable.[129] This is an idea borrowed from federal disability law.[130] The USERRA identifies undue hardship in the same way as the ADA defines the concept, that is, "actions requiring significant difficulty or expense."[131] Both statutes list the following factors to be considered when determining undue hardship: "the nature and cost of the action needed . . . the overall financial resources of the facility or facilities involved . . . the overall financial resources of the employer . . . and . . . the type of operation or operations of the employer."[132] The burden of proving undue hardship under both statutes is on the employer.[133] In both contexts, the gauge for determining undue hardship is to compare the financial costs associated with the reasonable efforts or accommodations in relation with the overall financial resources of, persons employed at, effect of expenses and resources, or the impact otherwise of the action on the facility's operation.[134]

There are ADA cases where the employees' absences from the job were similar to that under the USERRA—unscheduled and unpredictable. Specifically, the ADA has been applied to cases where the employees requested indefinite leave as a reasonable accommodation that would make them an otherwise qualified employee. In determining that it would be unreasonable to require employers to accommodate such absences, courts have focused not only on the frequency of these absences, but also on the burden to employers of

---

126. 508 F. Supp. 966 (E.D.N.C. 1981).
127. *Wrigglesworth*, 121 F. Supp. 2d at 1136 (citing *Davis*, 508 F. Supp. at 969).
128. Manson, *supra* note 2, at 74-75.
129. 38 U.S.C. § 4312(d)(1)(A), (B), (C) (2000).
130. 42 U.S.C. § 12112(b)(5)(A) (1994 & Supp. 1999).
131. 38 U.S.C. § 4303(15) (2000); 42 U.S.C. § 12111(10) (1994 & Supp. V 1999).
132. 38 U.S.C. § 4303(15) (2000); 42 U.S.C. § 12111(10) (1994 & Supp. V 1999).
133. 38 U.S.C. § 4312(d)(2) (2000); 42 U.S.C. § 12112(b)(5)(A) (1994 & Supp. V 1999).
134. 38 U.S.C. 4303(15)(B) (2000); 42 U.S.C. 12111(10)(B)(ii) (1994 & Supp. V 1999).

creating provisions at the last minute to cover for these absent employees.[135] Cases interpreting the federal statutes of the Rehabilitation Act and the ADA agree that chronic and excessive absenteeism need not be accommodated.

This section will first address the burden of proof in ADA cases that consist of the plaintiff proving an accommodation is reasonable and the employer proving that the accommodation would be an undue hardship. The next part of this discussion will explain the different ways indefinite leave is handled as an issue with respect to a reasonable accommodation in the case law. The final part of this section will discuss undue hardship with respect to ADA in case law. This section will explain the prevailing opinion that indefinite leave of absence is not "reasonable" because it does not enable a disabled person to work and the cost to any employer to train and pay a replacement worker to fill the same position or to continue operations short-handed for an indefinite period of time constitutes an undue burden on the employer. Thus, an indefinite leave of absence is not an "accommodation."

A. Unreasonable Versus Undue Hardship—the Burden of Proof

This subsection focuses on the burden of proof for ADA cases in which the disabled employee contends that indefinite leave is a reasonable accommodation and the employer contends that the accommodation is unreasonable or an undue hardship. Under the USERRA, after the initial elements required for reemployment have been met, the burden shifts to the employer to prove undue hardship.[136] By contrast, under the ADA, the disabled employee must meet the burden of proof that a reasonable accommodation exists before the burden shifts to the employer.

The Supreme Court recently agreed with holdings by lower courts that an employee need only show that an "accommodation" seems reasonable on its face.[137] The accommodation must be feasible for the employer and be objectively reasonable. An accommodation is reasonable only if its costs are not clearly disproportionate to the benefits it will produce.[138] The employee has the burden of production and can satisfy the burden of production by showing a plausible accommodation.[139]

After the employee makes the initial showing, the burden falls on the employer to show that the particular accommodation would be either unreasonable or cause the employer to suffer an undue hardship.[140] The employee makes his showing based on generality, while the employer must show special circumstances that demonstrate undue hardship in its particular

---

135. *See* Jackson v. Veterans Admin., 22 F.3d 277, 279 (11th Cir. 1994).

136. *See* 38 U.S.C. § 4312 (2003).

137. *See* U.S. Airways, Inc. v. Barnett, 535 U.S. 391, 401-02 (2002).

138. Borkowski v. Valley Cent. Sch. Dist., 63 F.3d 131, 138 (2d Cir. 1995).

139. *Id.*

140. *Id.*

circumstance.[141] The employer has far greater access to information than the typical employee, both about its own organization and about the practices and structure of the industry as a whole.[142] When the burden is passed to the employer, the difference between the burden of persuasion on the unreasonableness of the accommodation and demonstrating that the accommodation imposes an undue hardship becomes blurred.[143]

The Supreme Court's opinion in *Barnett* could be construed as ambiguous when attempting to reconcile the difference between the Second and Seventh Circuits.[144] However, the *Barnett* opinion could be read to suggest a practical way of reconciling the phrases "reasonable accommodation" and "undue hardship" and their burdens of proof. The Supreme Court explained that the terms do not mean the same thing:[145]

> a demand for an effective accommodation could prove *unreasonable* because of its impact, not on business operations, but on fellow employees—say because it will lead to dismissals, relocations, or modification of employee benefits to which an employer, looking at the matter from the perspective of the business itself, may be relatively indifferent.[146]

The undue hardship inquiry, on the other hand, focuses on the hardships imposed in the context of the particular employer's operations rather than the "general" reasonable standard for burden of proof required by the employee for reasonable accommodation.[147]

The burden on the employer to prove undue hardship consists of performing a cost/benefit analysis.[148] The employee meets his burden of production by identifying an accommodation that facially achieves a rough proportionality between costs and benefits.[149] However, an employer seeking to meet its burden of persuasion of reasonableness, accommodation, and undue hardship must undertake a more advanced analysis by arguing in the terms of § 12111(10).[150] In attempting to meet its burden of persuasion in establishing an affirmative defense of undue hardship, an employer does not have to analyze the costs and benefits of the proposed accommodation with "mathematical precision."[151] A common sense balancing of costs and benefits in light of the factors listed in the

---

141. *Barnett*, 535 U.S. at 401-02.
142. *Borkowski*, 63 F.3d at 137.
143. *Id.*
144. *Barnett*, 535 U.S. at 399-400.
145. *Id.*
146. *Id.* at 400-01 (emphasis added).
147. Barth v. Gelb, 2 F.3d 1180, 1187 (D.C. Cir. 1993).
148. *Borkowski*, 63 F.3d at 139.
149. *Id.*
150. *Id.*
151. *Id.* at 140.

statute and regulations is all that is needed.[152]

B. Indefinite Leave and Unreasonable Accommodation

1. Attending Work Regularly Is an Essential Function.—Some employers have attacked the issue of indefinite leave in ADA cases by arguing that being able to attend work regularly is an essential function of employee's job. An employer may argue that if an employee must go on indefinite leave, the employee cannot attend work regularly and, therefore, cannot perform the essential function of the job, thereby making the employee an unqualified individual. For instance, in *Tyndall v. National Education Centers Inc.*, the court held that because the employee's attendance problems rendered her unable to fulfill essential functions of her job and these problems occurred even with reasonable accommodations, she was not a qualified individual with a disability.[153] In that case, an employee worked for the employer in a teaching position, and while her performance was adequate, she was frequently absent due to her medical condition of lupus. In its holding, the court reasoned that an employee who does not come to work cannot perform any of her job functions, whether they are essential or otherwise.[154] Even an employee whose job performance is more than adequate when she is working will not be considered qualified for the job unless the employee is also willing and able to come to work on a regular basis.[155]

In *Nowak v. St. Rita High School*,[156] the court also held that regular attendance was an essential function of being a teacher. In *Nowak*, an employee had taught for over twenty-five years when he started having medical problems and missed sixty-five days. The employee worked ten days the next school year, and in the subsequent year the employee did not work at all. Because of the employee's extended illness and continual absence from the classroom, school administrators decided to terminate his employment. The court found that prior to his termination, Nowak was absent from his teaching position for more than eighteen months.[157] The court explained that "[t]he ADA does not require an employer to accommodate an employee who suffers a prolonged illness by allowing him an indefinite leave of absence."[158] Therefore, the court stated that Nowak was not a qualified individual with a disability because the federal law did not mandate that employers had to give leaves of absence to employees with

---

152. *Id.*

153. 31 F.3d 209, 213-14 (4th Cir. 1994).

154. *Id.* at 213.

155. *Id.*

156. 142 F.3d 999, 1003 (7th Cir. 1998).

157. *Id.* at 1003-04.

158. *Id.* (citing Rogers v. Int'l Marine Terminals, Inc., 87 F.3d 755, 759 (5th Cir.1996)) (stating "that nothing in the ADA requires an employer to accommodate an employee with an indefinite leave of absence and that because the plaintiff could not attend work, he was not a 'qualified individual with a disability' under the ADA").

prolonged illnesses.[159]

Under the USERRA, an employer does not have the option of arguing that the employee cannot do the job because of a lack of attendance and is therefore not qualified. The USERRA makes it impossible for an employee's military service to be a motivating factor in any termination of employment.[160] For an employer to argue against the statutory right of the employee to be reemployed, the employer can only argue the affirmative defense of undue hardship.[161] The employer cannot argue that missing work makes the employee unqualified. If the employee is not qualified to perform the duties in which he or she would have been employed when the employee returns from duty, the employer must make reasonable efforts to qualify the person for the new position.[162] Thus, the employer must provide retraining or upgrade training if the skills or technology are different when the employee returns from military duty. If those qualification efforts fail, then the employee must be returned to the position held on the date the military service began or "a position of like seniority, status and pay."[163]

2. Indefinite Leave Is Not a Reasonable Accommodation Under ADA.—If an employer unsuccessfully argues that indefinite leave makes an employee unqualified, then the employer can argue that indefinite leave is not a reasonable accommodation. In *Rogers v. International Marine Terminals, Inc.*, the court found that the employee was not a qualified employee under the ADA because he was not able to attend work at the time he was terminated, and the employer was not required to make a reasonable accommodation in the form of an indefinite leave of absence.[164] Rogers was employed as a mechanic who took paid sick leave for the treatment of persistent pain, swelling, and other problems in his right ankle. After using all his sick leave, Rogers received a year of disability benefits pursuant to a disability plan of the employer. The court found that Rogers remained unavailable from January 1993 to December 1993, and because Rogers could not attend work, he was not a "qualified individual with a disability" under the ADA.[165] The court agreed with several other courts that recognized "an essential element of any . . . job is an ability to appear for work . . . and to complete assigned tasks within a reasonable period of time."[166] Moreover, the court found that Rogers could not demonstrate that the employer could reasonably accommodate his purported disability.[167] In a similar case, the Fourth Circuit explained, "Nothing in the text of the reasonable accommodation provision requires an employer to wait an indefinite period for an

---

159. *Id.* at 1004.

160. *See* Klein & Cilenti, *supra* note 2, at 11.

161. *See id.* at 11-12.

162. *See* 38 U.S.C. § 4313(a)(1)(B) (2000).

163. *See id.* § 4313(a)(2)(B).

164. 87 F.3d 755 (5th Cir. 1996).

165. *Id.* at 759.

166. *Id.* (citing Carr v. Reno, 23 F.3d 525, 530 (D.C. Cir. 1994)).

167. *Id.*

accommodation to achieve its intended effect."[168] The court reasoned

> [a] reasonable accommodation is by its terms most logically construed as that which presently, or in the immediate future, enables the employee to perform the essential functions of the job in question. . . . [R]easonable accommodation does not require [an employer] to wait indefinitely for [the employee's] medical conditions to be corrected.[169]

In *Monette v. Electronic Data Systems Corp.*, the Sixth Circuit held that employers are "under no duty to keep employees on unpaid leave indefinitely until [a] position opens up."[170] In this case, Monette was a customer services representative for EDS who became injured delivering audio and visual equipment. Monette requested indefinite medical leave and received full pay and benefits for the next seven months. Monette proposed as a reasonable accommodation that the employer should have kept him on unpaid medical leave indefinitely until another customer service representative or receptionist position opened up. In holding for the employer, the court explained that employers simply are not required to keep an employee on staff indefinitely in the hope that some position may become available some time in the future.[171]

A few courts have found that an employee's request for an extended leave of absence could be a reasonable accommodation. For instance, in *Criado v. IBM Corp.*, the court upheld a jury verdict finding that IBM had violated the ADA by firing an employee who requested temporary leave to receive treatment for depression.[172] Holding that Criado's request for leave was a reasonable accommodation and did not unduly burden IBM,[173] the court recognized that a leave of absence for medical treatment may constitute a reasonable accommodation turning on the facts of the case.[174] The court found that the employer's unpaid leave policies were evidence that the employer would not be unduly burdened and could reasonably accommodate the employee's request.[175]

Another case in which a court found indefinite leave to be a reasonable accommodation is *Cehrs v. Northeast Ohio Alzheimer's Research Center.*[176] In *Cehrs*, an employee suffered from pustular psoriasis and psoriatic arthritis and could not work. The employer terminated her for allegedly failing to complete the paperwork to extend her leave of absence. The court held that an issue of fact existed as to whether granting medical leave would have unduly burdened the

---

168. Myers v. Hose, 50 F.3d 278, 283 (4th Cir. 1995).

169. *Id.*

170. 90 F.3d 1173, 1187 (6th Cir. 1996) (citing 42 U.S.C. § 12111(9)(A-B) (1994 & Supp. V 1999)).

171. *Id.*

172. 145 F.3d 437, 443 (1st Cir. 1998).

173. *Id.*

174. *Id.* at 443.

175. *Id.* at 444.

176. 155 F.3d 775 (6th Cir. 1998).

defendant or would have been a reasonable accommodation under the ADA.[177] The court contended that the employer never bears the burden of proving that the accommodation proposed by an employee is unreasonable, and it would impose an undue burden upon the employer if the a presumption existed that uninterrupted attendance is an essential job requirement.[178] The court explained that if an employer cannot show that an accommodation unduly burdens it, then there is no reason to deny the employee the accommodation.[179] It reasoned that Congress has already determined that uninterrupted attendance in the face of a family medical emergency is not a necessary job requirement and does not unduly burden employers.[180] Therefore, the *Cehrs* court concluded that there is no presumption that uninterrupted attendance is an essential job requirement, and it found that a medical leave of absence can constitute a reasonable accommodation under appropriate circumstances.[181] However, the court also held that the employee's Family and Medical Leave Act (FMLA) claim failed because the employee is allotted twelve weeks of absence by law, and the employee would have been unable to return within the twelve-week time period.[182]

The general rule in ADA cases concerning indefinite leave is that indefinite leave is not a reasonable accommodation because it would not enable the employee to perform the required essential functions. If an employee could not perform the essential functions, then the employee is not a qualified individual under the ADA. Some courts have extended Congress's intent of making unpaid leave a statutory right under the FMLA[183] to the ADA, allowing for employees to take unpaid leave as a reasonable accommodation.[184] However, courts can also read into the ADA the limitations Congress imposed on the FMLA of twelve weeks.[185] Therefore, the allowance for unpaid leave is not indefinite, but it is limited by some type of reasonableness standard.

For instance, in *Hudson v. MCI Telecommunications Corp.*,[186] an employee who worked as a customer service representative had Carpal Tunnel Syndrome. Her physician prohibited her from all typing and keyboard activity. The employee continued to work, after first being suspended for tardiness, until her termination. The court agreed with the plaintiff that a reasonable allowance of time for medical care and treatment may form a reasonable accommodation in appropriate circumstances.[187] However, the court found that the employee had

---

177. *Id.*
178. *Id.* at 782.
179. *Id.*
180. *Id.* at 783 (citing 29 U.S.C. §§ 2601-2654 (2000)).
181. *Id.*
182. *Id.* at 785 (citing 29 U.S.C. § 2612(a)(1)(D) (2000)).
183. Family and Medical Leave Act of 1993, 29 U.S.C. §§ 2601-2654 (1994 & Supp. V 1999).
184. *See* Criado v. IBM Corp., 145 F.3d 437 (1st Cir. 1998); *Cehrs*, 155 F.3d. at 775.
185. *Cehrs*, 155 F.3d. at 782.
186. 87 F.3d 1167 (10th Cir. 1996).
187. *Id.* at 1169.

failed to present any evidence of the expected duration of her impairment as of the date of her termination.[188] The court explained that MCI was not required to wait indefinitely for her recovery by either keeping her on its payroll or electing to pay the cost of her disability benefits.[189]

Likewise, in *Taylor v. Pepsi-Cola Co.*, the court affirmed a judgment that an indefinite period of medical leave was not a reasonable accommodation.[190] The court acknowledged that an allowance of time for medical care or treatment may constitute a reasonable accommodation.[191] However, the court cited *Rascon*[192] stating that an "indefinite unpaid leave is not a reasonable accommodation where the plaintiff fails to present evidence of expected duration of her impairment."[193]

Unlike the ADA, the USERRA has no provision enabling courts to check an employee's indefinite leave with a reasonable standard, except the length of five years. [194] Therefore, an employer must reemploy a returning serviceman regardless of how many times a serviceman leaves and how long the serviceman is gone, as long as five years of cumulative leave is not exceeded. An employer's only option under the USERRA is to carry the burden of proving undue hardship.

C. Undue Hardship on Employer

As first introduced, the ADA called for a very high threshold for undue hardship; an accommodation would not be unreasonable unless it threatened the continued existence of the employer's business.[195] Congress softened the burden on the employer by changing the definition to include a list of factors to be considered to determine what constitutes an undue hardship.[196] The most important factors in determining whether an accommodation causes undue hardship is "the employer's ability to bear the cost," rather than the cost of the accommodation itself.[197] Therefore, undue hardship must be determined on a case-by-case basis.[198] An accommodation that would impose an undue hardship "on a small business, or in a particular industry, may be reasonable for a large employer, or in a different industry."[199] Congress has clearly marked the limits

---

188. *Id.*

189. *Id.*

190. 196 F.3d 1106, 1110 (10th Cir. 1999).

191. *Id.*

192. Rascon v. U.S. West Communications, Inc., 143 F.3d 1324, 1333-34 (10th Cir. 1998).

193. *Taylor*, 196 F.3d at 1110.

194. *See* 38 U.S.C. § 4312 (2000).

195. Bonnie P. Tucker, *The Americans with Disabilities Act: An Overview*, 1989 U. ILL. L. REV. 923, 927 (1989).

196. *See* 42 U.S.C. § 12111(10)(A) (1994 & Supp. V 1999).

197. Jeffrey O. Cooper, *Overcoming Barriers to Employment: The Meaning of Reasonable Accommodation and Undue Hardship in the Americans with Disabilities Act*, 139 U. PA. L. REV. 1423, 1449 (1991).

198. *Id.*

199. *Id.* at 1450.

of undue hardship; "an undue hardship may be something less than a cost that would drive the employer to the verge of going out of business, but must impose more than a de minimis cost."[200] In other words, under the ADA, an accommodation imposes an undue hardship if its "cost would either . . . substantially impair the ability of the employer to produce goods or provide services, or . . . impose such a high cost that the employer would be forced to compensate by reducing the overall workforce."[201]

With respect to an employee taking indefinite leave as a reasonable accommodation, courts look to the specific facts of each case to determine whether an accommodation is an undue hardship on the employer. Courts have focused on the effect that an absence for an indefinite amount of time would have on other employees. For instance, in *Jackson v. Veterans Administration*, the court affirmed a judgment in favor of the employer and held that requiring the employer to accommodate the employee's repeated, sporadic, and unscheduled absences, caused because of a disability, by "making last-minute provisions for [the employee's] work to be done by someone else" would place undue hardship on employer.[202] The court rejected the proposed accommodation and stated that "[s]uch accommodations do not address the heart of the problem: the unpredictable nature of Jackson's absences. There is no way to accommodate this aspect of his absences."[203]

If an employee were able to take indefinite leave as an accommodation, the other employees would have to increase their workload in order for the employer to keep the employee's position open. The effect of increasing the workload on other employees is generally not considered a reasonable accommodation.[204] For instance, in *Turco v. Hoechst Celanese Corp.*, the court affirmed judgment in favor of an employer on an ADA claim where the employee's diabetes prevented him from meeting the physical or mental demands of his job.[205] The court rejected the contention that the employee would have been able to meet these demands better if he had been switched to a day shift and noted that because the employer did not have a day shift, the proposed accommodation would have imposed an undue burden on the employer by requiring that other employees work harder.[206]

To assert a defense of undue hardship, the employer must show more than increased costs of production. The employer must show that the increased cost threatens its ability to maintain its current level of output or its current workforce.[207] However, there is no bright line rule requiring quantitative evidence to prove undue hardship. There is only a general observation that

---

200. *Id.*
201. *Id.* at 1455.
202. 22 F.3d 277, 279 (11th Cir. 1994).
203. *Id.*
204. 29 C.F.R. § 1630.2 (2003).
205. 101 F.3d 1090, 1093 (5th Cir. 1996).
206. *Id.* at 1094.
207. Cooper, *supra* note 197, at 1456.

courts generally view evidence as less persuasive as it becomes more speculative and less quantitative.[208] In *Barth*, plaintiff, a computer specialist with a severe case of diabetes, wanted to work at one of twelve overseas radio relay stations, but his employer denied the application because he could not get medical clearance. Barth brought suit when his employer denied his request for a medical waiver and a limited waiver restricting his assignments to posts with suitable medical facilities.[209] The court held that the employer established sufficient facts to support a claim of undue hardship by virtue of the loss of "essential operational flexibility" that would have resulted from an attempt to accommodate Barth's medical needs.[210] The lower court observed that "the thin staffing at each post required flexibility of assignment, put a premium on workers not subject to serious health risks, and offered few options for initial assignment of Mr. Barth."[211]

As the case law may suggest, an employee on indefinite leave can subject an employee to undue hardship through its effect on the employer's operations as well as the effect on other employees. However, the employer still has the burden of proving undue hardship. In *Borkowski v. Valley Central School District*, the court held that the employer did not meet the burden of proof.[212] Ms. Borkowski, a school teacher for over twenty years, requested a teacher's aide to assist her in maintaining classroom control and thereby allow her to perform all of the functions of a library teacher. Ms. Borkowski presented some factual evidence that the accommodation would make her qualified, and that the accommodation was reasonable.[213] Therefore, the burden shifted to the employer, the school district. The court said the school district did not present any evidence concerning the cost of providing a teacher's aide, its budget and organization, or any of the other factors made relevant by the regulations.[214] The school argued that the provision of an assistant was unreasonable and created an undue hardship as a matter of law. However, the court concluded that there was nothing inherently unreasonable or undue in the burden that an employer would assume by providing an assistant to an employee with disabilities.[215]

> [I]n the absence of evidence regarding school district budgets, the cost of providing an aide of this sort, or any like kind of information, we are unable to conclude that unreasonableness or undue hardship has been established, and we certainly cannot say that either has been established as a matter of law.[216]

---

208. Barth v. Gelb, 2 F.3d. 1180, 1189 (D.C. Cir. 1993).
209. *Id.*
210. *Id.*
211. *Id.* at 1188.
212. 63 F.3d 131, 142 (2d Cir. 1995).
213. *Id.*
214. *Id.*
215. *Id.*
216. *Id.* at 143.

Therefore, employers generally must provide some type of statistical or quantified evidence from a cost and benefit analysis to meet their burden. However, in cases where indefinite leave for the employee is the issue, that is not necessarily the case. In *Walton v. Mental Health Association of Southeastern Pennsylvania*, Walton worked for "MHASP," an advocacy organization for people with mental illness.[217] Walton suffered from depression and because of her illness missed twenty-one days in 1990, forty days in 1991, fifty days in 1992, and fourteen and one-half days in 1993 before taking leave. MHASP terminated her a few months later.[218] The court affirmed that Walton's requested accommodation—continued leave—would have created an undue burden on MHASP.[219] The court stated that although unpaid leave supplementing regular sick and personal days might represent a reasonable accommodation, an employer does not have to allow this type of leave to the extent that MHASP had already granted it to Walton.[220] The court concluded that a blanket requirement that an employer allow such leave is beyond the scope of the ADA when the absent employee simply will not be performing the essential functions of her position.[221]

Though the USERRA and ADA have basically the same definition of undue hardship, the affirmative defense under the USERRA has been rarely used and almost never successful. There are several reasons for the difference in the successfulness of raising the affirmative defense under the two acts. First, under the ADA, employers are able to relate undue hardship to whether the employee can physically do the job and whether indefinite leave would allow the individual to do the job. However, an employee is generally considered unable to perform the essential functions if the individual is on an indefinite leave and not at work. Under the USERRA, undue hardship becomes merely a question of whether an employer can weather the cost of reemploying the returning serviceman after an indefinite amount of time. In interpreting the USERRA, courts presume that indefinite leave does not affect or threaten the essential functions of a job. Whereas, under the ADA, the judiciary has generally presumed indefinite leave has a significant effect on the essential functions of performing a job. Second, when the employer challenges an employee's claim under the ADA, generally the situation is reduced to the specific employer versus the specific employee claiming a disability and discrimination. Under the USERRA, however, the employer is perceived to be not only challenging the employee's reemployment, but also the nation's need for the citizen soldier to either defend its way of life or take time to ensure their readiness to defend its way of life. Therefore, an employer is less likely to publicly challenge an employee's right to reemployment. Third, related to the prior reason is the strength of an argument

---

217. 168 F.3d 661 (3rd Cir. 1999).

218. *Id.*

219. *Id.* at 671.

220. *Id.*

221. *Id.*

of reciprocal benefit under an employee leaving to serve his country that is not present under ADA cases.

IV. SOLUTION

A. Legislation

The first and most obvious solution to the problem of requiring employers to carry a disproportionate burden is legislation. The USERRA has a built-in presumption that indefinite leave is neither an unreasonable accommodation nor does it have any real effect on an employee's essential functions. A legislative amendment is needed to reduce the unfairness that the USERRA causes employers.

There has been some recent movement within Congress to decrease the burden employers are asked to bear. For instance, House Resolution 394 was introduced, which would allow employers a credit of up to $2000 for each Reservist who supports contingency operations in an active-duty status.[222] The legislation would give a $7500 total credit per employer.[223] As an owner of a Goodyear Tire Store explained, "In the past, companies have allowed their employees to serve the country because it's viewed as the patriotic thing to do. . . . Now, when two of my employees are fulfilling their military obligation, I've just lost one-third of my work force."[224]

Because of the economic slowdown and consequent fiscal tightening, House Resolution 394, along with other legislative efforts, has seen little legislative movement or success.[225] With legislative success seeming unlikely, the next best solution is for the judiciary to adopt the general approach of ADA burden-shifting with the added presumption that indefinite leave is unreasonable.

B. USERRA Burden Shifting

A solution to the lack of burden shifting element of the USERRA is the implementation of a similar burden shifting exercise employed by ADA case law. First, the burden shifting process should require the government to show that an indefinite leave is reasonable on its face, similar to the burden of the alleged disabled employee in ADA cases. The burden should be placed on the government because it is the government who the employee will generally work through to resolve reemployment situations.[226] This burden should be applied similarly to how the burden of showing indefinite leave was a reasonable

---

222. H.R. 394, 107th Cong. (2001).

223. *Id.*

224. Captain James R. Wilson, *Congress Considers Tax Breaks for Employers*, HILLTOP TIMES, Mar. 1, 2001, *at* http://www.hilltoptimes.com/archive/20010301/Supplement/2.html.

225. H.R. 394 (sent to House Comm. on Ways and Means); S. 3088, 107th Cong. (2002) (no co-sponsors and sent to Senate Comm. on Governmental Affairs); H.R. 3711, 107th Cong. (2002) (no co-sponsors and sent to House Comm. on Ways and Means).

226. *See* Klein & Cilenti, *supra* note 2, at 17.

accommodation in both *Rogers* and *Myers*.[227] In those cases, the courts held that indefinite leave was not a reasonable accommodation in the ADA claim because the employee could not meet the prima facie burden of showing that indefinite leave was generally a reasonable accommodation.[228]

During this process, the returning serviceman should be allowed to return to his position of reemployment. The issue is not whether the serviceman should be reemployed, but who should bear the cost of reemploying the serviceman. Therefore, if it is a situation where the reemployment would be found to be an unreasonable accommodation, then the remedy should be for the government to reimburse the employer for all training and other expenses involved in the reemployment of the returning serviceman. For instance, for an airline pilot who leaves for twelve months of duty and goes through the minimum retraining required, costs average slightly over $10,000 for the employer.[229] If the period of leave is found to be unreasonable, the government would reimburse the airline for the retraining expense. In this type of situation, the facts would generally not allow for a finding that an airline that loses one or even 200 out of 8000 pilots is faced with an undue hardship. Yet for an airline to be faced with bearing the costs of retraining its pilots only to have them called to duty again would be unreasonable.

If the government can show that the period of leave was reasonable, then the burden should shift to the employer to show that reemployment would be an undue hardship. For undue hardship, the burden is not necessarily based on statistical proof. Beyond when reemployment may be unreasonable, undue hardship should concern when the indefinite leave of a serviceman could actually put the operations of a business in jeopardy. Under the USERRA, the courts liberally construe the statutes in favor of protecting the serviceman. If the end is not the reemployment of the serviceman, but a determination of who should bear the cost, the courts should either construe the statute evenly for both parties or in favor of the employer. For situations where the employer is able to prove undue hardship, the government should provide compensation to the employer for keeping positions open for an indefinite amount of time. In addition, for a small employer with less than fifty employees, finances should be provided to the employer for the cost of keeping the position open.

The question should not be whether the employee should be reemployed because that could decrease volunteerism and force the country to rely on involuntary measures. Involuntary measures consist of what the military has most recently used over the last couple of years such as a constant "stop-loss" policy. "Stop-loss" is a policy which holds servicemen who wish to retire or leave the military involuntarily, or a measure which has not been used since the Vietnam era—conscription. The question should be who bears the cost of

---

227. *See* Rogers v. Int'l Marine Terminals, Inc., 87 F.3d. 755 (5th Cir. 1996); Myers v. Hose, 50 F.3d 278 (4th Cir. 1995).

228. *Id.*

229. Interview with David Mikkelson, First Officer on Airbus 328, United Airlines, in Chicago, Ill. (Jan. 16, 2003).

reemploying the returning servicemen.

C. Economic Argument

The military's move to a "Total Force Policy" and greater reliance on the citizen soldier has not been an issue unique to our time. The issue of whether our nation's military should consist of a professional standing army or a group of militias has been around since the colonial years. Adam Smith believed that the ability of a nation to wage war is best measured in terms of its productive capacity.[230] Military theorists in Smith's time felt that security demanded a well-trained and well-disciplined armed force to battle their adversaries. It was generally thought that a militia, however trained and disciplined, could not take the place of professional soldiers, especially in an age when the development of firearms put a greater premium on organization and order than on individual skill, bravery, and dexterity.[231]

Like Adam Smith, Alexander Hamilton also believed that the professional army should be the basis of a national defense, not a band of militias. Hamilton wrote, "[W]ar, like most other things, is a science to be acquired and perfected by diligence, by perseverance, by time, and by practice."[232] On the other hand, Thomas Jefferson saw a professional standing army as being a possible threat to the individual liberties of its own people.[233] Though he did agree with Hamilton that the military establishment and some type of constant force was necessary for the nation's defense, Jefferson thought a system of universal liability where all men participate in the nation's military was more suitable than a professional standing army.[234]

The military has moved to a "Total Force Policy" to increase the numbers in the military at a decreased cost.[235] In 1990, the total active duty force was 2.065 million servicemen and the military budget was approximately $390 billion, or 5.8% of the Gross Domestic Product (GDP).[236] In 2002, the total active duty force was less than 1.4 million servicemen, the military budget was

---

230. MAKERS OF MODERN STRATEGY FROM MACHIAVELLI TO THE NUCLEAR AGE 222 (Peter Paret et al. eds., 1986).

231. *Id.* at 228.

232. *Id.* at 239 (citing THE FEDERALIST, No. 25, 156 (1787)).

233. *Id.* at 241.

234. *Id.* at 242. In a letter to James Monroe in 1813, Jefferson wrote: "[I]t proves more forcibly the necessity of obliging every citizen to be a soldier; this was the case with the Greeks and Romans, and must be that of every free State .... We must train and classify the whole of our male citizens and make military instruction a regular part of collegiate education. We can never be safe till this is done."

235. *See* Fernandez, *supra* note 5, at 861.

236. Department of Defense, *News Transcript, Background Briefing on the Fiscal Year 2002 Budget Amendment, at* http://www.natprior.org/charts/DOD1977to2002.htm (last modified June 22, 2001).

approximately $300 billion, which is only 3.7% of GDP.[237] As the military transitioned out of the cold war and began decreasing the size of its force, the leadership realized that a Reserve unit could be maintained in a mission-ready status similar to an active duty unit at one third the cost. Therefore, a military force structure could be maintained relying heavily on the citizen soldier and still allow for large reductions in the military budget. The civilian leadership has focused on the large savings in a citizen soldier army without realizing the true social costs. This is because a large portion of the cost has been externalized to other parties, predominantly the employer of the citizen soldier.

Externalities exist whenever some person or entity makes a decision about how to use resources without taking full account of the effects of the decision. The entity ignores some of the effects—some of the costs or benefits that result from a particular activity because they fall on others.[238] As a consequence of this external cost shift, resources tend to be misused or "misallocated."[239] In this case, the government is placing the burden on the employer by shifting the cost. The military is using the citizen soldier similarly to a private company relying on temporary workers, thereby relieving itself of the burden of having to pay for the benefits generally due full time employees. The employers, on the other hand, have to maintain the benefits of the citizen-soldier upon the individual's return from duty as if he never left. The term "externalities" implies that when external costs are found, steps should be taken by the government to eliminate them.[240]

One could argue that the benefit the citizen soldiers create in defending their country and its way of life is felt by everyone. Therefore, the externalities are reciprocal because the employer benefits from security and stability, by-products of the citizen-solder's efforts. However, the cost is disproportionately placed on the employer. Most citizens have to pay taxes in one form or another, a portion of which goes to pay for the national defense. Employers not only pay their share of taxes, but also involuntarily have their employees taken from them often and for indefinite periods of time. Such employers are then burdened with maintaining the citizen-soldier's employment position and reemploying the returning serviceman in a position that he would have been in as if he had never left. The cost that the government saves in not having to maintain a large professional force has been shifted disproportionately to the employer. Pigou writes: "In any industry, where there is reason to believe that the free play of self-interest will cause an amount of resources to be invested different from the amount that is required in the best interest of the national dividend, there is a prima facie case for public intervention."[241] The aim of economic policy is to ensure that people, when deciding which course of action to take, choose that which brings about the best outcome for the system as a whole.[242]

---

237. *Id.*

238. JESSE DUKEMINIER & JAMES E. KRIER, PROPERTY 47 (4th ed. 1998).

239. *Id.*

240. RONALD HARRY COASE, THE FIRM, THE MARKET, AND THE LAW 27 (1988).

241. A.C. PIGOU, THE ECONOMICS OF WELFARE 5th ed. (1952).

242. COASE, *supra* note 240, at 27.

The way in which industry, the military in this instance, is organized is thus dependent on the relation between the costs of carrying out transactions in the market and the costs of organizing the same operations within that firm which can perform the task at the lowest cost.[243] The lowest cost for the military to organize its operations is to rely on the citizen soldier. The military in effect will increase its allocation of resources and organize its operations to rely more heavily on the citizen soldier rather than on the professional soldier until the marginal costs of relying on the citizen-soldier are greater than the cost of the professional soldier. The producer, the military, is normally only interested in maximizing its own incomes or in this case, minimizing its costs and is not concerned with social costs.[244] The military will only undertake an activity if the cost of the factors employed is less than its private cost.[245]

The solution is not to expect the government to negotiate with each employer of service personnel because the transaction costs would be far too high. Furthermore, there is the practical reciprocal nature of the problem, in that the military must have the discretion to use citizen soldiers when the need arises for the country's national security. However, expecting the military to negotiate with employers or allowing the country's security to be vulnerable by creating obstacles to the military relying on the citizen soldier is not an option. Instead, the burden shifting adapted by the courts in USERRA cases should be made more similar to that of the ADA cases as discussed above. When under a new burden of proof scheme, an employer can show that reemployment is unreasonable or an undue burden. The government should then provide compensation to the employer to ease the reemployment of the serviceman. In that way, the military can still use the citizen solder, and the employer has a realistic opportunity to meet the burden of showing reemployment is unreasonable or an undue burden.

D. FMLA Comparisons

Indefinite leave is not necessarily always an unreasonable accommodation. With the creation of the Family and Medical Leave Act (FMLA) and the trend in some ADA cases since the creation of the FMLA, indefinite leave has become more of an accepted accommodation.[246] Under the FMLA, an employee becomes eligible for FMLA coverage if he or she has been employed by a covered employer for no less than a year and has worked at least 1250 hours during the preceding twelve months.[247] Once eligible, an employee may take reasonable periods of unpaid leave for medical reasons, for child-birth or adoption, or for the care of a spouse, parent, or child who suffers from a serious health condition.[248] Leave periods are circumscribed: an eligible employee may take a maximum of

---

243. *Id.* at 63.
244. *Id.* at 158.
245. *Id.*
246. Cehrs v. Northeast Ohio Alzheimer's Research Ctr., 155 F.3d 775 (6th Cir. 1998).
247. 29 U.S.C. § 2611(2)A (2000).
248. *Id.* § 2601(b)2.

twelve workweeks of FMLA leave in any twelve-month span.[249] Following such a leave, an employee is entitled to reclaim his or her former job.[250]

However, the FMLA applies only to private sector operations that employ fifty or more persons.[251] Applying the FMLA to smaller employers would unfairly burden a small employer by causing it to go without an employee for as much as a three-month period once a year. Under the USERRA, reemployment rights apply to all employers regardless of size. Furthermore, the FMLA implicates shorter time frames, which, as defined, can be an illness that lasts as little as four days.[252] Added to this, the maximum annual benefit under the FMLA is twelve weeks of unpaid leave.[253] While leave periods under the USERRA are many times truly indefinite, the periods of leave traditionally last between one weekend and twenty-four months. [254]

In *Jordan*, the court found that the citizen-soldier had an unqualified right to reemployment as long as the duration requirements were met.[255] An accumulative duration requirement of five years of leave is not a reasonable accommodation. The cost to an employer to have to retrain employees who leave for six, twelve, or twenty-four month periods is an unfair and disproportionate burden. Under the FMLA, an employer knows that the employee will be returning no later than three months from the date she left. Under the USERRA there is no way to know the duration of the leave except that it cannot be longer than five years.

To avoid the transactions costs of having to engage in a burden-shifting process like the ADA, perhaps an approach similar to the FMLA would be appropriate. For periods of less than three months of leave, the government would reimburse employers only for retraining expenses of returning serviceman. For periods over three months, then allow typical USERRA case law approach of requiring the employer to prove undue hardship to recover additional expenses such as of having to hire and train third parties to temporarily replace the departed citizen soldier, or business losses directly related to having the employee leave.

CONCLUSION

The USERRA was created to protect the citizen soldiers' right to reemployment in a time when Congress could not have foreseen or anticipated the demand and the resultant reliance on the citizen soldier. This Note has attempted to show the unfairness of the reemployment provision in the USERRA.

---

249. *Id.* § 2612(a)1.

250. *Id.* § 2614(a)1.

251. *Id.* § 2611(4).

252. *Id.* 2612(a)(1)(C) (stating that an employee may qualify for FMLA leave to care for a child under eighteen merely by showing that the child suffers from a serious health condition).

253. *Id.* 2612(a)(1) (2000).

254. Reuters, *U.S. Reserves May Stay on Duty for Second Year*, *available at* http://www.ledger-enquirer.com/mld/ledgerenquirer/3942819.htm (last visited Aug. 26, 2002).

255. Jordan v. Air Prods. & Chems., Inc., 225 F. Supp. 2d 1206 (C.D. Cal 2002).

There have been several efforts to correct this within Congress and give employers some type of financial help to reduce the impact of having to reemploy employees who are gone for indefinite periods of time. Because of the political and economic realities of our time, the unfairness will have to be corrected by the judiciary. The judiciary can find an equitable solution by adapting a burden of proof similar to that used in ADA cases. The judiciary could also use an equity approach that attempts to take into consideration the externalized costs that the reemployment right places on the employer. The question is not whether citizen soldiers should have the right to reemployment, but rather, who should bear the ultimate cost in fulfilling that right.

A SAFER NATION?: HOW DRIVER'S LICENSE RESTRICTIONS HURT IMMIGRANTS & NONCITIZENS, NOT TERRORISTS

ALEXANDER L. MOUNTS*

INTRODUCTION

After the tragic events of September 11, 2001, federal and state government officials called for sweeping reforms of the nation's and states' security laws.[1] This reform has been ongoing and has impacted the daily lives of people living in the United States by affecting everything from increasing luggage screening in airports to reviewing the validity of visas held by noncitizens. In this process of reform to help prevent terrorism, accessibility to a driver's license has been a targeted issue which has lead to heated discussions from many different groups representing many different points of view.[2] Why driver's licenses? Apparently, several of the terrorists involved in the September 11 attacks had United States state-issued driver's licenses that allowed them to board the airplanes.[3] However, what officials failed to mention was that none of the hijackers needed driver's licenses to accomplish their goal; all of the hijackers had foreign passports that served as valid identification at airports.[4] While increasing national security is critical in the current world *milieu*, restricting driver's licenses is an ineffective way to enforce immigration laws and prevent terrorism.

An additional policy justification officials give for imposing increased restrictions on driver's licenses is to protect the public from identity theft.[5] In

---

* J.D. Candidate, 2004, Indiana University School of Law—Indianapolis; B.A., 2001, Taylor University, Upland, Indiana. I would especially like to thank my wife Whitney who continually inspires me to strive for excellence with her unconditional love, support and encouragement.

1. During the 2001-2002 state legislative sessions, sixty-three bills were introduced addressing the ability of the immigrants to obtain a driver's license. *Immigrant Driver's License Restrictions Challenged in Some States*, NATIONAL IMMIGRATION LAW CENTER, Vol. 16, No. 6 (Oct. 21, 2002), *available at* http://www.nilc.org/immspbs/ DLs/DL004.htm.

2. *Immigrant Driver's License Proposals and Campaigns: Surprising Progress Since 9/11*, NATIONAL IMMIGRATION LAW CENTER (May 14, 2002), *at* http://www.nilc.org/immspbs/DLs/DL002.htm.

3. Albert Harberson & Eileen Doherty, *Driver's License Integrity*, THE COUNCIL OF STATE GOVERNMENTS AND THE NATIONAL CONFERENCE OF STATE LEGISLATURES (June 25, 2002), *available at* http://www.aamva.org/documents/idsNCSL_CSGWhitepaperToHomelandSecurity.pdf; *Driver's License Fact Sheet: Protecting the Public Safety and National Security of Everyone—Driver's Licenses, National Security, and Terrorism*, NATIONAL COUNCIL OF LA RAZA, *available at* http://nclr.policy.net/proactive/newsroom/release.vtml?id=19942 [hereinafter *Driver's License Fact Sheet: Protecting the Public Safety and National Security of Everyone*].

4. *Drivers License Fact Sheet: Protecting the Public Safety and National Security of Everyone*, *supra* note 3.

5. Identity theft occurs when someone uses the "identifying information [of another person]—name, social security number, mother's maiden name, or other personal information—to commit fraud or engage in other unlawful activities." *Prepared Statement of the Federal Trade*

2000, an estimated 500,000 to 700,000 Americans were victims of identity theft.[6] With the advent of the Internet and ever-increasing technology, thieves are constantly finding new ways to steal personal information and fraudulently use it to access accounts or credit lines. Many states assert that requiring additional forms of identification when applying for a license will prevent identity theft by making people prove their identities.[7] This policy justification assumes that requiring more papers, not better verification, is the solution to preventing identity theft.

Increased driver's licenses restrictions, as passed and as previously existing, do not apply exclusively to a handful of terrorists living in sleeper cells in this country. According to the 2000 United States Census, there are over thirty million immigrants in the United States, representing 11% of the total population.[8] Between 1970 and 2000, the naturalized citizen population increased by 71%.[9] It is axiomatic that the immigrant population is not only substantial, but is also on the rise. Despite the myth that immigrants are a drain on the United States economy, in the year 2000, the foreign-born population accounted for 12.4% of the total civilian labor force.[10] These immigrants collectively earned $240 billion a year, paid $90 billion a year in taxes and received $5 billion in welfare.[11] Offsetting the amount immigrants contributed in taxes, the use of public benefits by legal immigrant families with children who earn less that 200% of the federal poverty level fell sharply between 1994 and 1999.[12]

Commission on Identity Theft Before the Subcommittee on Technology, Terrorism and Government Information of the Committee on the Judiciary, U.S. Senate (statement of Jodie Bernstein, who was at the time Director, Bureau of Consumer Protection), *available at* http://www.ftc.gov/os/2000/03/identitytheft.htm.

6. *Identity Theft: How It Happens, Its Impact on Victims, and Legislative Solutions*, Hearing on Pub. Law 105-318 Before the S. Subcomm. on Tech., Terrorism, and Gov't Info. of the Subcomm. on the Judiciary, 106th Cong. 31-32 (statement of Beth Givens, Director, Privacy Rights Clearinghouse), *available at* http://www.privacyrights.org/ar/id_theft.htm. *But see* Harberson & Doherty, *supra* note 3, at 6. Harberson and Doherty feel that the 500,000 figure is inflated by as much as 500% due to the way fraud was reported by consumers. The General Accounting Office has since changed the way it collects data, which has reduced the instances of "identity theft" from 500,000 cases to approximately 100,000 cases per year.

7. *See* Erin Gartner, *Driver's License Now Harder to Get; To Deter Terrorism and I.D. Theft, BMV Tells Newcomers to Furnish More Proof*, INDIANAPOLIS STAR, July 13, 2002, at A1, *available at* 2002 WL 2390485.

8. Michael Fix et al., *The Integration of Immigrant Families in the United States*, URB. INST. (July 2001), *available at* http://www.urban.org/Uploaded PDF/immig_integration.pdf.

9. United States Census Bureau, *Profile of the Foreign-Born Population in the United States: 2000*, U.S. DEPARTMENT OF COMMERCE, Dec. 2001.

10. *Id.*

11. *Immigration Myths and Facts: Five Immigrations Myths Explained*: AMERICAN IMMIGRATION LAWYERS ASSOCIATION, *at* http://www.aila.org/contentViewer.aspx?bc=17,142 (copyright 2003).

12. Michael Fix & Jeffrey Passel, *The Scope and Impact of Welfare Reform's Immigrant*

Part I of this Note addresses the foundation and evolution of driver's license restrictions in the United States. Part II of this Note analyzes two current justifications given by states for the restrictions. The first of these justifications is maintaining the integrity of license issuance: ensuring that the person receiving a license meets necessary driving competency standards and has a verifiable identity. The second justification is prevention of identity theft: providing appropriate law enforcement officials the ability to verify the authenticity of the license document, driving history and identity of the license holder. Part III addresses two constitutional concerns that have arisen with the increased restrictions. The first potential concern is an Equal Protection Clause dilemma, while the second potential concern is a Supremacy Clause dilemma. Finally, after weighing the strengths and weaknesses of the justifications and constitutional concerns, Part IV of this Note addresses the public policy reasons why increased restrictions for driver's licenses are an ineffective way to prevent identity fraud and terrorism.

I. DRIVER'S LICENSE HISTORY

A. The Foundation

The issuance of driver's licenses has always been a function of the individual states. The authority for control over the issuance of driver's licenses is derived from the Tenth Amendment to the U.S. Constitution. The Amendment reads, "[t]he powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people."[13] Historically, the federal government has not threatened to disturb the states' power in this area; however, the federal government has limited the states' power in the area of commercial driver's licenses (CDLs).[14] CDLs represent approximately 5% of the total number of licenses issued in the United States.[15] The Commercial Motor Vehicle Act of 1986 (CMVSA) established the Commercial Driver's License Program, which requires states to ensure that drivers with certain serious driving violations be prohibited from operating a Commercial Motor Vehicle (CMV).[16] Should the states not comply with federal government requirements, sanctions can be imposed.[17]

The first state-driven driver's license law was passed by Rhode Island in

---

Provisions, URB. INST. 16 (Jan. 2002), *available at* http://www.urban.org/uploadedPDF/410412_discussion02-03.pdf.

13. U.S. CONST. amend. X.

14. 49 U.S.C. § 31102 (2003).

15. Harberson & Doherty, *supra* note 3, at 2. There are 190 million licensed drivers in the United States.

16. Memorandum from the Texas State Office of Risk Mgmt. to the State Risk Managers (Sept. 4, 2002) *at* http://www.sorm.state.tx.us/Notices/FedDriverLicStandards090402.htm.

17. *Id.*

1908.[18] From that time to the present, every territory and state has passed statutory provisions and administrative regulations to govern the privilege of driving.[19] Standards were immediately developed to protect public safety by recognizing those individuals who met the prescribed standards. These standards generally include a minimum age requirement, a physical ability requirement, a practical driving competency requirement and a requirement for knowledge of traffic laws.[20]

B. The Evolution

The modern use of the driver's license has expanded to include ancillary uses such as cashing checks, obtaining library cards and boarding airplanes; however, the primary purpose of the license has remained unchanged: to provide state sanctioning in order to operate a vehicle. As previously mentioned, the states' power of control over the license has remained relatively unscathed. One reason for the lack of involvement of the federal government can be attributed to the proactivity of the states.

The states have traditionally had the primary responsibility for building and regulating the nation's highways. Because a driver's license allows the public to move easily between states, two interstate compacts have been formed. These contractual agreements among the states are the Driver License Compact (DLC) and the Nonresident Violator Compact (NRVC).[21] The DLC, created in 1961, ensures that a driver's home state receives and processes information about traffic violations committed by the driver in another state. A total of forty-five states have adopted the DLC.[22] The NRVC standardizes methods used by different jurisdictions to process traffic citations received by out of state drivers. Both compacts are administered by the American Association of Motor Vehicle Administrators (AAMVA).[23] While the states have enjoyed a large degree of freedom in their control over driver's licenses, Congress is currently considering three pieces of legislation that would "federalize" the driver's license, thus, preempting states' control over the issuance of driver's licenses.[24]

---

18. Harberson & Doherty, *supra* note 3, at 1.

19. Since 1954, all states have required drivers to be licensed. *Id.* at 2.

20. *Id.* at 1.

21. *See* THE COUNCIL OF STATE GOVERNMENTS, THE DRIVER'S LICENSE COMPACT AND THE VEHICLE EQUIPMENT SAFETY COMPACT, IN INTERSTATE COMPACTS FOR TRAFFIC SAFETY (1962).

22. Harberson & Doherty, *supra* note 3, at 2.

23. *Id.*

24. The Commerce Clause and spending power of Congress would likely not preclude federal activity in the area of licensing individuals. The Commerce Clause would likely apply because our roads and highways are crucial to commerce and travel between states. In a Commerce Clause analysis, the court must defer to congressional findings that the regulated activity substantially affects interstate commerce, as long as there is any rational basis for such finding and, given such finding, the only remaining question for judicial inquiry is whether there is reasonable connection between the regulatory means selected and the asserted ends. *See* U.S. CONST. art. I, § 8, cl. 3; Houston, E & W.T.R. Co. v. United States, 234 U.S. 342 (1914) (appeal from the United States

While it is axiomatic that no state licensing agency would voluntarily issue a license to a terrorist or identity thief, it is impossible to imagine an issuance system that could contemplate every intended use for the license in an effort to screen out some individuals. For example, licensing agencies would not have been able to predict the case of Lucas Helder, the clean-cut, twenty-one year old college student, who admitted to planting eighteen pipe bombs in five states during a weekend cross-country spree.[25] Some have even suggested that "federalizing" the sale of box cutters would be a more relevant response to the September 11 attacks than "federalizing" the driver's license.[26]

The first piece of proposed federal legislation is the Driver's License Modernization Act of 2002 (DLMA), introduced by Representatives Jim Moran and Tom Davis of Virginia.[27] The proposed legislation would require driver's licenses to become "smart cards" with computer chips. Biometric data would be collected to match the license with the owner. States' participation would be required. Tamper-resistant security features would be incorporated into all license documents. Lastly, states would be required to adopt and implement procedures for accurately documenting the identity and residence of an individual before issuing a license.[28]

The second piece of legislation, the Driver's License Integrity Act of 2002, is similar to the Driver's License Modernization Act. Illinois Senator Richard Durbin's legislation would require minimum uniform standards for issuance and administration of state-issued driver's licenses, sharing of driving information between the states for verification, with enhanced privacy protection, enhanced capabilities for authentication and verification of licenses, penalties for internal fraud, and similar state funding allocation.[29] The final piece of legislation was introduced by Representative Jeff Flake of Arizona and would bar any federal agency from accepting a state-issued driver's license as a valid identification unless the state requires licenses issued to nonimmigrant noncitizens to expire upon the expiration of the alien's nonimmigrant visa.[30]

It is unlikely that a federally mandated solution will emerge because a

---

Commerce Court to review judgments dismissing the petitions in suits to set aside an order of the Interstate Commerce Commission regulating railway rates). Congress' spending power could be used to condition funds, such as those used for highways, on the state's willingness to comply with the federal guidelines set forth. *See* South Dakota v. Dole, 483 U.S. 203, 211-12 (1987) (federal statute conditioning states' receipt of portion of federal highway funds on adoption of minimum drinking age of twenty-one).

25. Charles Feldman et al., *Feds: Suspect Admitted Pipe Bomb Spree*, CNN.COM (May 9, 2002), *available at* http://www.cnn.com/2002/US/05/08/mailbox.pipebombs/. It is interesting to note that Lucas Helder was stopped three times by the police during the weekend and only the last police officer realized he was driving with an expired license. Helder was issued a ticket and allowed to go on his way.

26. Harberson & Doherty, *supra* note 3, at 5.

27. *Id.* at 10.

28. *Id.* at 10-11.

29. *Id.* at 11.

30. *Id.*

significant number of weaknesses exist. One weakness is that any federal legislation would potentially be subject to years of constitutional legal challenges and might be struck down.[31] A second weakness is one of time constraints. It took the federal government six years to implement the commercial driver's license after passing the enacting legislation.[32] Moreover, Daniel Hartman, manager of the division over CDLs at the United States Department of Transportation, stated in May 2002 that the most significant problem that system faces after sixteen years in existence is fraud in the licensing authority offices.[33] A third weakness is that federal legislation requires the setting of standards without making funds available to pay the costs.[34]

Now that a holistic examination has been performed on the origin of driver's licenses, who has control over the licenses, from where the power to control the licenses came and what attempts on control have been performed by the states and by the federal government, a solid foundation exists to begin to analyze the legitimacy of the current driver's licenses restrictions imposed by states in reaction to September 11, 2001.

II. CURRENT JUSTIFICATIONS FOR THE RESTRICTIONS

"For many individuals, a Driver['s] License . . . issued by the Indiana Bureau of Motor Vehicles (BMV) is the most important means of proving their identity. The Bureau of Motor Vehicles endeavors to safeguard the integrity of driver documents and to protect the public from false and/or fraudulent applications."[35] This statement came from the Commissioner of the Indiana Bureau of Motor Vehicles. Indiana's response is typical of other states in its current approach to licensing standards. In fact, the preceding quotation was made in response to revised requirements the state was implementing on September 30, 2002. States provide two main justifications in response to questions concerning new restrictions placed on licenses. The first justification is the need to maintain integrity of license issuance, and the second justification is the need to prevent identity theft.

A. Maintain Integrity of License Issuance

The states' first justification for restrictions on driver's licenses, which is really a disguise for the prevention of terrorism, is the need to maintain integrity of license issuance by ensuring that the person receiving a license meets necessary driving competency standards and has a verifiable identity. The competency standards requirement has no bearing on the increased restrictions because no new competency requirement was added. These standards, as

---

31. *Id.* at 12.

32. *Id.*

33. *Id.*

34. *Id.*

35. INDIANA BUREAU OF MOTOR VEHICLES, *Driver License Identification Requirements*, *at* http://www.in.gov/bmv/driverlicense/idreq.html (last visited Feb. 18, 2003).

previously mentioned in Part I of this Note, include meeting the proper age requirement, passing a standardized test and meeting certain physical requirements.[36] Each state varies in its specific competency requirements, but the aforementioned core requirements usually remain unchanged.[37] Because nothing regarding this component has been changed, opponents to the new driver's license restrictions have not made mention of any problems with this requirement.

However, the second portion of maintaining the integrity of license issuance has been the subject of much controversy because it is the primary area affected by the increased restrictions. States that imposed increased restrictions pertaining to whether a person has a verifiable identity did so in three areas: a social security number requirement, a proof of identity requirement, and a proof of residency requirement/legal immigration status requirement. Each of the three areas will be explored in depth to discover how each is related to the purported justification given.

1. The Social Security Requirement.—The Social Security Number (SSN) was created under the Social Security Act and was originally designed to keep track of an individual's earnings and eligibility benefits. It is now used by both governmental and nongovernmental entities for numerous purposes.[38]

Social Security Requirements for Each State [39]

States that require a SSN for a driver's license (DL) without exceptions (5):	Alabama, District of Columbia, Hawaii, New Jersey and West Virginia
States that do not require a SSN for a DL (7):	Georgia, Kansas, Maryland, Minnesota, Mississippi, Oregon and Vermont
States that require a SSN for a DL only of people who have been assigned one or are eligible for one (34):	Alaska, Arizona, Arkansas, California, Colorado, Delaware, Florida, Idaho, Indiana, Louisiana, Maine, Massachusetts, Missouri, Montana, Nebraska, Nevada, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Washington, Wisconsin and Wyoming
States that require a SSN but have other exceptions to the rule (5):	Connecticut, New Hampshire, Illinois, Kentucky and Iowa

36. Harberson & Doherty, *supra* note 3, at 1.

37. *Id.*

38. *Driver's License Fact Sheet: Ensuring Immigrant Access to Driver's Licenses: Alternatives to Social Security Number Requirements*, NATIONAL COUNCIL OF LA RAZA, *at* http://www.nclr.policy.net/proactive/newsroom/release.vtml?id=19940 (last visited Nov. 14, 2002) [hereinafter *Driver's License Fact Sheet: Ensuring Immigrant Access to Driver's Licenses*].

39. *Driver's Licenses for Immigrants: Broad Diversity Characterizes States' Requirements*, NATIONAL IMMIGRATION LAW CENTER (Nov. 6, 2002), *at* http://www.nilc.org/immspbs/DLs/DL005.htm (last visited Oct. 1, 2003).

Of the thirty-four states that only require a SSN from people who have been assigned one or are eligible for one, some allow for the submission of an affidavit stating that they are ineligible for one or have never received one.[40] Others require submission from the Social Security Administration that no SSN has been assigned or that the application for one has been denied.[41] A few states allow applicants without a SSN to submit an individual taxpayer identification number. Lastly, three of the states require applicants without SSNs to verify lawful presence.[42] Prior to March 1, 2002, the Social Security Agency would assign numbers to lawful residents who did not have work authorization but still needed a SSN for non-work related reasons, such as acquiring a driver's license.[43] Since March 1, 2002, the Social Security Agency no longer assigns SSNs when the sole reason for issuing a SSN is to obtain a driver's license. Some people who are lawfully present in the United States but who are not authorized to work are not able to gain access to a SSN, which is a prerequisite for a driver's license in some states.

In 1996, two pieces of federal legislation were introduced that addressed SSN use for the purpose of obtaining a state-issued driver's license. The first was the Personal Responsibility and Work Opportunity Reconciliation Act of 1996 (PRWORA), otherwise known as "Welfare Reform."[44] The second was the Illegal Immigrant Reform and Immigrant Responsibility Act of 1996 (IIRIRA), otherwise known as "Immigrant Reform."[45]

The "Welfare Reform" Act contained a provision requesting corresponding state agencies in charge of driver's license issuances to record the SSN of applicants for certain licenses.[46] The Act also requested SSNs be recorded for medical documents and certain court-issued documents for child support enforcement purpose. Section 466(a)(13)(A) of the Act required a SSN for any commercial driver's license applications.[47] Subsequent legislation was enacted applying the SSN requirement for all driver's licenses, not just commercial driver's licenses.[48]

Implementation of the expanded requirement of Section 466(a)(13)(A) was difficult for Departments of Public Safety and Motor Vehicles because they were uncertain of the proper relationship between the expanded requirement and accessibility to a driver's license.[49] The staff persons and advocates requested

---

40. *Id.*

41. *Id.*

42. *Id.*

43. Michele L. Waslin, Ph.D., *Safe Roads, Safe Communities: Immigrants and State Driver's License Requirements*, NATIONAL COUNCIL OF LA RAZA, No. 6, May 2003, at 3, *available at* http://www.nclr.org/policy/briefs (last visited Nov. 21, 2002).

44. *Driver's License Fact Sheet: Ensuring Immigrant Access to Driver's Licenses*, *supra* note 38, at 1.

45. *Id.*

46. *Id.*

47. *See id.*

48. *Id.* at 1-2.

49. *Id.* at 2.

a proper interpretation from the appropriate governing agency, the United States Department of Health and Human Services, because the section dealt with child support enforcement.[50] In response, Commissioner David Gray Ross of the Department of Health and Human Services offered this interpretation: "'We interpret Section 466(a)(13)(A) to require that States have procedures, which require an individual to furnish any Social Security Number that he or she may have. [However,] Section 466(a)(13)(A) does not require that an individual have a social security number as a condition of receiving a [driver's] license.'"[51] Whether state licensing agencies were aware of the proper interpretation of Section 466(A)(13)(A) or not, many state driver's licensing administrators currently require a SSN as a condition to apply for a license.[52]

The second piece of legislation, "Immigrant Reform," required all state driver's licensing agencies to request the SSNs of all driver's license applicants and place the SSN on the driver's license.[53] The effect of this legislation was to create a *de facto* national identification card. Various advocacy groups joined with states in opposition to this provision on the grounds that it would violate privacy rights, lead to increased discrimination against immigrants and certain ethnic groups and lead to an increase in identity theft.[54] Section 656(b) of IIRIRA of 1996 was repealed in October 1999 due to overwhelming opposition.[55]

2. The Proof of Identity Requirement.—A second heightened restriction aimed at maintaining the integrity of license issuance is the addition or strengthening of a proof of identity requirement. Every state has a list of redundant documents that it accepts to verify the identity of noncitizens. However, many individuals, who are lawfully in the United States, can be excluded at various stages of the immigration process. Some immigrants are unable to produce the required documentation to prove their identity because they do not have any, or enough, of the acceptable documents. Most states have similar requirements for the documents required for proper verification to exist. Generally, a combination of what a state deems as a primary and a secondary document is needed.[56] Some examples of acceptable documents include: U.S. passports, U.S. original state birth certificates, driver's licenses issued by other countries, student identification cards, original Social Security Cards, tribal photo ID, United States military photo identification card, some Immigration and Naturalization Services (INS) documents (Certificate of Naturalization, an Arrival-Departure Record (I-94), an Alien Registration Receipt Card (I-551)), a

---

50. *Id.*
51. *Id.*
52. *Id.* at 2.
53. *Id.* at 1.
54. *Id.*
55. *Id.*
56. Foundation documents, or primary documents, are used to provide the building blocks of personal information on which the license is issued. These documents range from birth certificates, to passports, to other state driver's licenses. Secondary documents, although they vary from state to state, include W-2 tax forms, a school report card, a consulate-issued identification card or a major credit or bank card.

letter of authorization issued by the INS, a Visa, or a Valid Employment Authorization Card (I-688 A or B) or a Matricula Consular.[57]

A Matricula Consular is an identity document issued by the Mexican Consulate; currently thirteen states accept this as a valid form of identification.[58] Before a Mexican citizen can obtain a Matricula Consular, he or she must present a certified copy of his or her birth certificate and picture identification to the Consulate.[59] This type of document is gaining recognition in the United States.[60] On November 7, 2001, the Consulate General of Mexico and Wells Fargo announced that the Matricula Consular is an acceptable form of primary identification for opening new accounts and over-the-counter transactions at its more than 3000 banking locations in twenty-three states.[61] Bank One and Fifth Third Bank followed suit in September 2002.[62]

3. The Proof of Residency/Legal Immigration Status Requirement.—States vary in the last area of maintaining integrity of license issuance, the proof of residency and legal immigration status requirement. California, for example, explicitly requires proof of legal immigration status or proof of legal residency in the United States.[63] Other states are not as explicit. The South Carolina Department of Motor Vehicles interprets its state statute more narrowly by not granting driver's licenses to immigrants without green cards or valid student visas.[64]

---

57. Waslin, *supra* note 43, at 3-4.

58. NATIONAL IMMIGRATION LAW CENTER, *supra* note 39. States that accept the Matricula Consular as a valid form of identification (13): Idaho, Indiana, Michigan (accepted on a case-by-case basis), Nebraska, North Carolina, New Mexico, Oregon, South Dakota, Tennessee, Texas, Utah, Washington and Wisconsin.

59. Waslin, *supra* note 43, at 12.

60. One such example is Chicago's adoption of the document. *See* Cook County, IL, Ordinance Recognizing Matricula Consular as Legal Identification, Section 5-4 (Sept. 19, 2002).

61. Press Release, Wells Fargo, Wells Fargo To Accept Matricula Consular Card As Identification For New Account Openings: Consulate General of Mexico and Wells Fargo Make Announcement in Los Angeles (Nov. 9, 2001), *at* http://www.wellsfargo.com/press/article.jhtml?path=eis%2Fpress% 2Fmatricula20011109b.jhtml&year=2001.

62. John C. Ashe, *Bank One, Fifth Third Bank Now Accept Mexico ID Card, Open Doors to New Business*, MICHIGAN BANKER, Sept. 11, 2002, *available at* http://www.michiganbankermag.com/14691.htm.

63. Waslin, *supra* note 43, at 4.

64. *Id.* South Carolina Code section 56-1-40(7) (2002) denies a license to anyone who is not a resident, but includes in its definition of a resident "all persons authorized by the United States Department of Justice, the United States Immigration and Naturalization Service, or the United States Department of State to live, work, or study in the United States on a temporary or permanent basis who present documents indicating their intent to live, work, or study in South Carolina."

States' Lawful Presence Requirements [65]

States that have lawful presence requirements (27):	Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, District of Columbia, Florida, Georgia, Idaho, Iowa, Kansas, Kentucky, Louisiana, Maine, Minnesota, Mississippi, Nebraska, New Jersey, Oklahoma, Pennsylvania, South Carolina, South Dakota, Virginia, West Virginia and Wyoming
States that do not have lawful presence requirements (24):	Alaska, Hawaii, Illinois, Maryland, Massachusetts, Michigan, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Mexico, New York, North Carolina, North Dakota, Ohio, Oregon, Rhode Island, Tennessee, Texas, Utah, Vermont, Washington and Wisconsin

As a result of states' impositions of increased requirements to protect the maintenance of the license issuance (meeting necessary driving competency standards and having verifiable identity: social security requirement, proof of residency requirement and legal immigration status) many legal immigrants without proper documentation are currently being denied access to driver's licenses because they cannot meet the increased requirements.

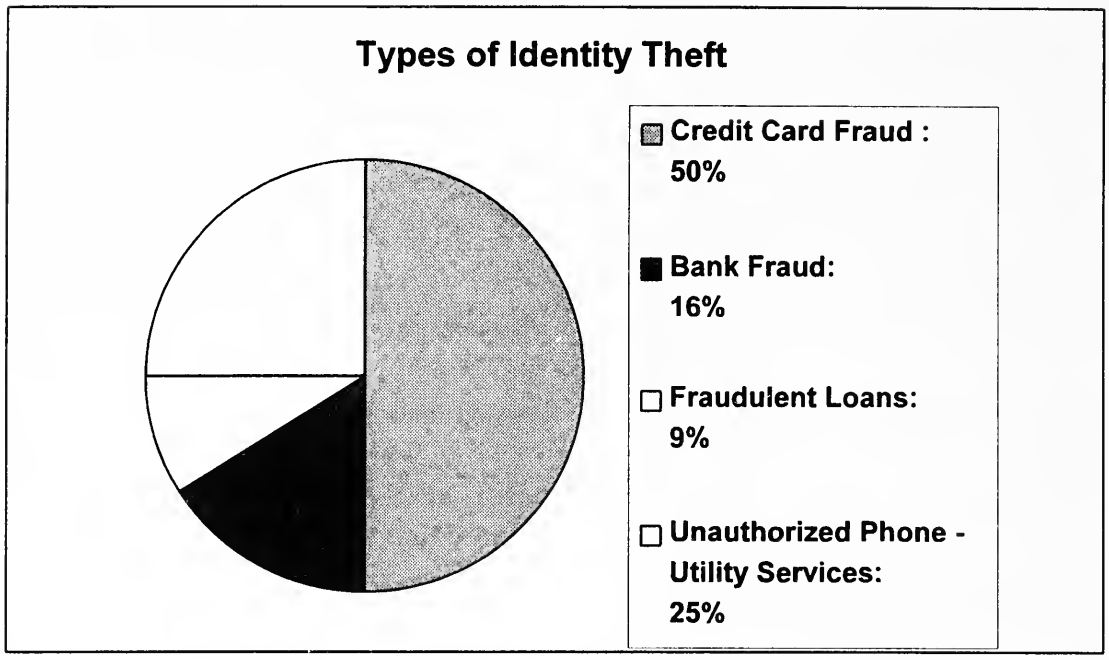
B. Prevention of Identity Theft

The second justification for restrictions on driver's license issuance is the need to prevent identity theft by providing appropriate state law enforcement officials the ability to verify the authenticity of the license document, the driving history and the identity of the license holder. In analyzing this justification, it is first important to look at data concerning the growing trend of identity theft in the United States. The Federal Trade Commission provides a break down of compiled statistics collected nationwide on identity theft.[66]

---

65. NATIONAL IMMIGRATION LAW CENTER, *supra* note 39, at 2.

66. *Identity Theft Complaint Data: Figures and Trends on Identity Theft January 2000 through December 2000*, IDENTITY THEFT DATA CLEARING HOUSE, FEDERAL TRADE COMMISSION 2, *available at* http://www.ftc.gov/bcp/workshops/idtheft/trends-update_2000.pdf.



Of the percentages listed on the chart, only eight percent (8%) of the victims reporting identity theft indicated that his or her identity had been obtained from a government document (e.g., a forged driver's license, a fraudulently filed document, such as a tax return, or a conversion of his or her government benefits).[67] Of the persons indicating identity fraud in this way, only .22% f the cases were reported due to fake driver's licenses.[68] A small percentage of the reported identity theft from driver's licenses is not trivial, and states should continue to seek ways to protect their citizens from different sources of potential harm, including driver's license fraud. Nevertheless, the question must be asked whether states' efforts to fortify this weakness are really effective.

While many states believed that requiring more documents would help to prevent identity theft, a growing dissent has formed among activist groups. The principal challenge related to foundation documents is states' ability to verify their authenticity and validity. Dr. Richard Varn, Chief Information Officer for the State of Iowa,[69] recently expounded on this very issue in his testimony before the Senate Subcommittee on Oversight of Government Management:

> Identity security is a critical component of ensuring accuracy, preventing fraud, and granting privileges and benefits in many programs and processes . . . . The identity system [referring to driver's license system] . . . is broken and is more likely to actually enable identity theft and fraud

67. *Id.*

68. Harberson & Doherty, *supra* note 3, at 6.

69. NASCIO STATE PROFILES IOWA, *available at* http://www.nascio.org/aboutNascio/profiles/iowa.cfm (last visited Jan. 24, 2003). Dr. Varn is responsible for information technology operations and policy for the state of Iowa and works directly for Iowa Governor Tom Vilsack.

> rather than prevent it . . . . As a result, facts such as social security number, address, birth day, and mother's maiden name . . . can be used to create identity and extend privileges and benefits fraudulently. It is not these facts or our inability to keep them secret that is the problem: it is that we rely on them alone to establish identity.
>
> . . . .
>
> [O]ur driver identity systems, card, systems, security measures, and issuance processes are not uniformly and adequately conducted and coordinated to ensure transportation safety let alone the myriad of their other uses on which we have come to depend.[70]

Only a handful of states actively verify foundation documents; however, even if states do attempt to verify the documents, it is nearly impossible to rely on state employees' ability to recognize an authentic foundation document. An extreme example of this is the 16,000 different U.S. birth certificates.[71]

Linda Foley, director of the Identity Theft Resource Center based in San Diego, California, believes the problem with driver's licenses is that they have become universally accepted as positive identification.

> [T]he social security card and birth certificates—do not directly prove the holder is the person of record. Birth certificates only verify birth records. Both are easily forged or purchased by anyone. Both are public records and not protected nor private. Neither definitely link through photo or fingerprints with the cardholder. I think you see the problem. We have a positive identifier [foundation document] that is built on sand—nothing solid or confirmable.[72]

Department of Motor Vehicle offices throughout the country are prime targets for those who commit identity fraud and theft. *Dateline NBC*, a news program, featured an undercover investigation to demonstrate the ease of fraudulently obtaining a driver's license.[73] The program highlighted how "brokers" provide customers with fake SSNs and walk them through the licensing process.[74] The investigative report participant was able to fraudulently obtain a

---

70. *Test. before the S. Subcomm. on Oversight of Gov't Mgmt., Restructuring and the D.C.: A Hearing Regarding: A License to Break the Law? Protecting the Integrity of Driver's Licenses*, Apr. 16, 2002, *available at* http://www.itd.state.ia.us/Varn_presentations/testimony_narrative.htm (presented by Richard J. Varn, CIO, State of Iowa, on behalf of the Nat'l Ass'n of State Chief Info. Officers (NASCIO), the Nat'l Gov's Ass'n (NGA), and the Info. Tech. Dept. (ITD), State of Iowa).

71. Harberson & Doherty, *supra* note 3, at 3.

72. *Test. for Cal. Leg. S. Comm. On Transp.: Special Oversight Hr'g on Identity Theft—Cal. Dep't of Motor Vehicles* (Nov. 16, 2000) (presented by Linda Foley, director of the Identity Theft Resource Center).

73. *Dateline NBC: Dateline Hidden Camera Investigation: Identity Crisis: How Easily False Driver's Licenses Can Be Obtained* (NBC television broadcast, Oct. 26, 2001), *available at* WL24017777.

74. *Id.*

California driver's license, even though the SSN belonged to a Florida doctor.[75] California is not the only state suffering from problems with integrity of license issuance. North Carolina estimates that 388,000 people in its motor vehicle system have given 999-99-9999 as their SSN.[76]

Other states have faced internal fraud problems from their employees. The *Star-Ledger* of Newark, New Jersey reported in the summer of 2002 that eight New Jersey Department of Motor Vehicle employees were being indicted as part of various rings selling fake driver's licenses.[77] In response to this internal fraud problem, Eric Shuffler, Chief of Staff for the New Jersey State Transportation Commissioner, said, "New Jersey's driver's license is, however, one of the easiest in the nation to counterfeit, and document fraud is a rampant problem in the state."[78] The *Orange County Register* reported in 2000 that the California Department of Motor Vehicle employees were selling driver's licenses for up to $4000 each and that sixty active cases of fraud existed.[79] With the recent discovery of internal fraud among state departments, it is difficult to determine just how many people lost their identity, not due to immigrants fraudulently giving documents, but rather by a government worker making money on the side issuing fraudulent documents.

III. CONSTITUTIONAL PROBLEMS ARISING FROM DRIVER'S LICENSE RESTRICTIONS

Over the past century, three principles have been formed by judicial decisions regarding immigration, federalism, and equal protection for immigrants. The first of these principles is that the federal power, in regard to immigration, is plenary and not bound by foreign affairs or national security and is immune from judicial scrutiny.[80] The second principle is that noncitizens fall within "persons" protected by the Equal Protection Clause of the Fourteenth Amendment.[81] The

---

75. *Id.*

76. Tracey Lackman, Comment, *Welfare Reform Conflict: An Analysis of 42 U.S.C.A. § 666(A)(13)(a) of The Personal Responsibility and Work Opportunity Reconciliation Act of 1996*, 19 T.M. COOLEY. L. REV. 105, 123 (2002).

77. Robert Schwaneberg, *State Vows to End DMV Neglect-User Friendly Services and Tighter Security Via Digital Licenses in Works*, STAR-LEDGER, Aug. 8, 2000, at 2, *available at* 2002 WL 25387972.

78. Brian Donohue & Mark Mueller, *Hijackers Lose Spot in DMV Sale Pitch—Administration Admits Lack of Fraud Proof*, STAR-LEDGER, Dec. 15, 2002, at 1, *available at* 2002 WL 103818305.

79. Harberson & Doherty, *supra* note 3, at 4.

80. Michael J. Wishnie, *Laboratories of Bigotry?: Devolution of the Immigration Power, Equal Protection, and Federalism*, 76 N.Y.U. L. REV. 493, 501 (May 2001).

81. *Id.* The pertinent part of the Fourteenth Amendment reads, "nor deny to any person within its jurisdiction the equal protection of the laws." U.S. CONST. amend. XIV. *See* Yick Wo v. Hopkins, 118 U.S. 356, 369 (1886) (provisions of the Fourteenth Amendment "are universal in their application, to all persons within the territorial jurisdiction, without regard to any differences of race, of color, or of nationality; and the equal protection of the laws is a pledge of the protection

last principle is that the federal government grants states no similar plenary power in regard to regulating or overseeing immigration.[82] This part of the note will focus on the last two principles.

A. Equal Protection Problem

Immigrants may bring Fourteenth Amendment equal protection challenges to discriminatory state actions.[83] Many of these discriminatory actions, such as the imposition of greater restrictions on obtaining a driver's license, knowingly or unknowingly, cause discrimination between two different groups of people: American citizens and certain classes of noncitizens. "[L]egal immigrants may invoke the Fourteenth Amendment's Equal Protection Clause against discriminatory state measures, and the plenary power doctrine[84]does not shield states from more searching scrutiny."[85]

In agreement, the U.S. Supreme Court in *Graham v. Richardson*[86] found that permanent resident noncitizens are a "discrete and insular minority."[87] In *Graham*, Justice Blackmun, for the Supreme Court, held that provisions of state welfare laws conditioning benefits on citizenship and imposing durational residency requirements on noncitizens were violative of the Equal Protection Clause of the Fourteenth Amendment.[88] Because this group of persons has

---

of equal laws"). *Id.* The Court went on to hold that San Francisco had violated the Fourteenth Amendment by engaging in impermissible discrimination based on "hostility to the race and nationality" of Chinese immigrants. *Id.* at 374. *See also* Torao Takahashi v. Fish & Game Comm'n, 334 U.S. 410, 427 (1948).

> We should not blink at the fact that § 990 [the statute in question], as now written, is a discriminatory piece of legislation having no relation whatever to any constitutionally cognizable interest of California. It was drawn against a background of racial and economic tension . . . . We need but unbutton the seemingly innocent words of § 990 to discover beneath them the very negation of all the ideals of the Equal Protection Clause.

Id. at 427. *See also* Truax v. Raich, 239 U.S. 33, 39 (1915). For a more complete analysis of the United States Supreme Court's equal protection analysis with regard to immigrants, see *Developments in the Law—Immigration Policy and the Rights of Aliens*, 96 HARV. L. REV. 1286, 1299-1302 (1983).

82. Wishnie, *supra* note 80.

83. *Id.*

84. The plenary power doctrine, read broadly, is interpreted to mean that exercises of federal immigration power are bound up in national security and foreign affairs. This interpretation allows the federal government to be largely immune from searching judicial review. *Id.*; see also Linda Bosniak, *Membership, Equality, and the Difference that Alienage Makes*, 69 N.Y.U. L. REV. 1047 (1994) for a thoughtful analysis of the plenary power doctrine.

85. Wishnie, *supra* note 80, at 501 (footnote added).

86. 403 U.S. 365 (1971).

87. *Id.* at 371-72 (quoting United States v. Carolene Prods. Co., 304 U.S. 114, 152 n.4 (1938)).

88. *Id.* at 376.

historically been discriminated against and unable to protect itself from this discrimination, courts have closely scrutinized any state discrimination against legal immigrants and have frequently invalidated it.[89]

Because permanent residents (noncitizens and certain legal immigrants) have important civic responsibilities, such as paying taxes and subjection to military conscription, on an equal basis with citizens, laws that treat permanent residents differently should come under close scrutiny.[90] In undertaking this scrutiny, "the governmental interest claimed to justify the discrimination is to be carefully examined in order to determine whether that interest is legitimate and substantial, and inquiry must be made whether the means adopted to achieve the goal are necessary and precisely drawn."[91] With driver's license restrictions, the question the courts should ask is whether increased restrictions on driver's licenses are "necessary and precisely drawn" to prevent terrorism and identity theft in the United States.

In beginning an equal protection analysis, a court must first determine whether the new restriction discriminates between two groups.[92] Advocates in support of increased restrictions could argue that two distinct groups do not exist because not all immigrants or noncitizens are hurt by the statute; therefore, they argue, the statute does not discriminate against the class. However, the court in *Nyquist* said, "[t]he important points are that . . . [the 'discriminatory' statute being challenged] is directed at aliens and that only aliens are harmed by it. The fact that the statute is not an absolute bar does not mean that it does not discriminate against the class."[93] Today, some revised statutes do not even mention immigrants or noncitizens. Nevertheless, the practical effect of the revised statutes coupled with the policy justifications mentioned in Part II of this Note and given by department of motor vehicle branches are that the revised statutes are equally aimed at discriminating between citizens and lawfully present immigrants and noncitizens.[94] When these factors are viewed holistically, the effect of the statutes is such that the revision might as well have been implemented with immigrants and noncitizens in mind.

The second step in the equal protection analysis requires undertaking whether the governmental interest justifying the discrimination is substantial and

---

89. Wishnie, *supra* note 80, at 502; *see, e.g.*, Nyquist v. Mauclet, 432 U.S. 1, 7-8 (1977); *Graham*, 403 U.S. at 365.

90. *Graham*, 403 U.S. at 376; Wishnie, *supra* note 80, at 505.

91. *Nyquist*, 432 U.S. at 7 (quoting Examining Bd. v. Flores de Otero, 426 U.S. 572, 605 (1976)). It is interesting to note that in *Nyquist*, the New York statute in question barred both loans and scholarships for higher education to some groups of noncitizens while allowing certain other groups of noncitizens to receive the scholarships. This is comparable to allowing some noncitizens or immigrants to possess a driver's license while disallowing others. *See also* Sugarman v. Dougall, 413 U.S. 634 (1973) (striking down a civil service law provision that only citizens could be eligible for certain levels of civil service employment).

92. *See* Plyler v. Doe, 457 U.S. 202, 216-17 (1982).

93. *Nyquist*, 432 U.S. at 9.

94. An important reminder is that the majority of these restrictions have been proposed in response to the September 11 attacks.

legitimate.[95]

In viewing the two policy justifications for increased driver's license restrictions given by states, prevention of terrorism and identity theft, both facially appear to serve a legitimate and substantial interest. The ultimate state interest in licensing drivers is "the important public interest in safety on the roads and highways, and in the prompt removal of a safety hazard."[96] If departments of motor vehicles are attempting to achieve their policy justifications by ensuring identity, it is unclear why noncitizens, who are in the country lawfully, meaning that this citizen has documentation that satisfies the federal government's requirements, are potentially more of a hazard to license than a United States citizen.[97] The rule is not narrowly tailored to meet the state's interest in assuring safety on the roads.

If states, by way of increased restrictions, are attempting to verify identification for either the prevention of terrorism or identity theft, they may be performing a function reserved for the federal government. "Control over immigration and naturalization is entrusted exclusively to the Federal Government, and a State has no power to interfere."[98]

The previously cited cases in this section all involved state-invoked discrimination against noncitizens who were legally in the United States. In these cases, the courts applied strict or close scrutiny. Had the discrimination existed on a federal level, the discrimination would have likely been upheld in accordance with plenary power doctrine that would invoke a "narrow standard of review of decisions made by the Congress or the President in the area of immigration . . . ."[99] The double standard between the state and federal application has been a point of contention with critics who argue that all non-citizen classifications should be treated under a single standard.[100]

Nonetheless, some noncitizens or immigrants who are unlawfully present in the United States are not members of a suspect class.[101] Discrimination against these persons will be analyzed by the standard "that the classification at issue bears some fair relationship to a legitimate public purpose."[102] This is an easier

---

95. *Nyquist*, 432 U.S. at 7 (quoting *Examining Bd.*, 426 U.S. at 605).

96. Dixon v. Love, 431 U.S. 105, 114 (1977).

97. Taking Indiana's safety interest as an example, Indiana Code section 9-25-4-1 (2002) only requires that potential drivers pass the prescribed test and that the potential drivers obtain insurance.

98. *Nyquist*, 432 U.S. at 10.

99. Matthews v. Diaz, 426 U.S. 67, 82 (1976).

100. Wishnie, *supra* note 80, at 508.

101. *See* Plyler v. Doe, 457 U.S. 202, 219 n.19 (1982) (a Texas statute which withheld from local school districts any state funds for education of children who were not "legally admitted" into the United States and which authorized local school districts to deny enrollment to such children). The Court noted that "[u]ndocumented aliens cannot be treated as a suspect class . . . ." *Id.* at 223. The Court then acknowledged that education is not a fundamental right. *Id.* Nevertheless, the Court recognized that the Texas statute imposed "a lifetime hardship on a discreet class of children not accountable for their disabling status." *Id.*

102. *Id.* at 216.

standard for states to meet in an effort to defend their actions. A judge or jury would probably not find it difficult to prove that a state government's restrictions have a fair relationship to a legitimate state-interest in verifying the identity of a potential driver's license holder. In fact, in 2001 a judge found such an interest in *Doe v. Georgia Department of Public Safety*.[103] The ruling upheld the constitutionality of denying licenses to illegal immigrants and emphasized the existence of a legitimate interest.[104]

It is possible, but not likely, that a court could find that a legitimate relationship does not exist between restricting driver's licenses to noncitizens here unlawfully and state interests in assuring safety and protecting identity. Noncitizens could unlawfully, if provided the chance, take the required written and driving tests and obtain insurance. If allowed to do so, the increased restrictions would not rationally relate to the safety objective because it would not be clear why noncitizens would pose a greater hazard than citizens. However, to be successful in arguing that the discrimination does not bear a fair relationship to a legitimate state interest,[105] a judge or jury would also have to be convinced that proof of identification bears no relation to citizenship status. This argument would likely fail because, as previously mentioned, a lesser standard of review is applied when reviewing cases involving illegal immigrants.

B. The Supremacy Clause Problem and Preemption by Congress

The second potential constitutional problem that might be encountered with the imposition of driver's license restrictions relates to the principle that the federal government grants states no similar plenary power in regard to regulating or overseeing immigration.[106] The Supremacy Clause of the United States Constitution reads:

> This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United Sates, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in

---

103. 147 F. Supp. 2d 1369 (11th Cir. 2001).

104. *Id.* at 1376. The court listed three reasons for finding a legitimate purpose: because Georgia has a legitimate interest in limiting its services to citizens and legal residents; second, Georgia has a legitimate interest in not allowing its governmental machinery to be a facilitator for the concealment of illegal noncitizens; and third, Georgia has a legitimate interest in restricting Georgia driver's licenses to those who are citizens or legal residents because of the concern that persons subject to immediate deportation will not be financially responsible for property damage or personal injury due to automobile accidents. *Id.*

105. *Plyer*, 457 U.S. at 216.

106. Wishnie, *supra* note 80, at 501. *See also* F.M.C. Corp. v. Holliday, 498 U.S. 52, 56-7 (1990). When looking to see if a federal law pre-empts a state statute, the courts look to congressional intent. "Pre-emption may be either express or implied, and is compelled whether Congress' command is explicitly stated in the statute's language or implicitly contained in its structure and purpose." *Id.*

the Constitution or Laws of any State to the Contrary notwithstanding.[107]

Michael J. Wishnie, Assistant Professor of Clinical Law at New York University, commented on immigration and federalism in his article, *Laboratories of Bigotry? Devolution of the Immigration Power, Equal Protection, and Federalism.* He noted:

> In immigration . . . the federal government has reigned supreme for over a century. Even before the federal government expanded its own regulation of immigration in the 1880s, the Supreme Court invalidated state and local efforts to regulate immigration or legal immigrants when those measures conflicted, expressly or implicitly, with federal immigration policy.[108]

It would appear that any state regulation of noncitizens was an ipso facto regulation of immigration and was thus preempted; however, this is not the case. In *DeCanas v. Bica*,[109] the Court unanimously noted otherwise, pointing out that

> [e]ven if such local regulation has some purely speculative and indirect impact on immigration, it does not thereby become a constitutionally proscribed regulation of immigration that Congress itself would be powerless to authorize or approve. Thus, absent congressional action [the local statute] . . . would not be an invalid state incursion on federal power.[110]

Although states are free to regulate in some situations where Congress has not, there are times when the state regulation, even though harmonious with federal regulation, must give way to paramount federal legislation. In *Michigan Canners & Freezers Ass'n v. Agricultural Marketing & Bargaining Board*,[111] the Court repeated that three ways exist in which federal law may pre-empt state law.[112]

> First, in enacting the federal law, Congress may explicitly define the extent to which it intends to pre-empt state law.[113] Second, even in the absence of express pre-emptive language, Congress may indicate an

---

107. U.S. CONST. art. VI. *See also* U.S. CONST. art. I, § 8, cl. 4 ("The Congress shall have the Power . . . *To establish an uniform Rule of Naturalization* . . . .").

108. Wishnie, *supra* note 80, at 510 (internal citations omitted).

109. 424 U.S. 351 (1976) (8-0 decision) (Stevens, J., not participating).

110. *Id.* at 355-56 (holding that the California Labor Code provision which prohibited unlawfully present persons from being knowingly employed if the employment would injure lawfully present workers was not preempted under the Supremacy Clause as a regulation of immigration).

111. 467 U.S. 461 (1984) (holding that the extent the Michigan Agricultural Marketing and Bargaining Act conflicted with the Federal Agricultural Fair Practices Act by establishing "accredited" associations which wielded the power to coerce producers to sell their products according to terms established by the association and to force producers to pay a service fee for the privilege, it was preempted by the federal act).

112. *Id.* at 469.

113. *Id.*; *see also* Shaw v. Delta Air Lines, Inc., 463 U.S. 85, 95-96 (1983).

> intent to occupy an entire field of regulation, in which case the States must leave all regulatory activity in that area to the Federal Government.[114] Finally, if Congress has not displaced state regulation entirely, it may nonetheless pre-empt state law to the extent that the state law actually conflicts with federal law. Such a conflict arises when compliance with both state and federal law is impossible, . . . or when the state law "stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress."[115]

1. Congress May Explicitly Define the Extent to Which It Intends to Preempt State Law.—In this case the applicable federal laws would be the Immigration and Nationality Act (INA).[116] Because the federal government has historically refrained from involvement with state-issued driver's licenses, it is not likely that a court would find that the Constitution or the INA explicitly preempts state action in this area.[117]

2. Congress May Intend to Occupy an Entire Field of Regulation.—Using the second way to show preemption, even in the absence of express preemptive language, Congress may intend to occupy an entire field of regulation since

> [T]he Federal Government has broad constitutional powers in determining what aliens shall be admitted to the United States, the period they may remain, regulation of their conduct before naturalization, and the terms and conditions of their naturalization. Under the Constitution the states are granted no such powers; they can neither add to nor take from the conditions lawfully imposed by Congress upon admission, naturalization and residence of aliens in the United States or the several states. State laws which impose discriminatory burdens upon the entrance or residence of aliens lawfully within the United States conflict with this constitutionally derived federal power to regulate immigration, and have accordingly been held invalid.[118]

Even if the state law is a permissible regulation of immigration, it may still be preempted if there is a showing that it was the "clear and manifest purpose of Congress" to effect a "complete ouster of state power, including state power to promulgate laws not in conflict with federal laws" with respect to the subject

---

114. *Mich. Canners & Freezers Ass'n*, 467 U.S. at 469 (citations omitted); *see also* Fid. Fed. Sav. & Loan Ass'n v. De la Cuesta, 458 U.S. 141, 153 (1982).

115. *Mich. Canners & Freezers Ass'n*, 467 U.S. at 469; *see also* Fla. Lime & Avocado Growers, Inc. v. Paul, 373 U.S. 132, 142-43 (1963); Hines v. Davidowitz, 312 U.S. 52 (1941).

116. Immigration and Nationality Act, 8 U.S.C.A. § 1101 (2003).

117. *See supra* text accompanying note 24. This is not to say, however, that the federal government could not use the Commerce Clause and the Spending Power to interfere with state control of the driver licensing system. Should such an event take place and states had primary control, then preemption might take place.

118. Takahashi v. Fish & Game Comm'n, 334 U.S. 410, 419 (1948) (internal citations omitted); *see also Hines*, 312 U.S. at 66.

matter which the statute attempts to regulate.[119]

Advocates of increased driver's license restrictions argue that even though it is apparent that Congress intends to occupy the entire field of immigration, imposing increased restrictions on driver's licenses does not bring such an action within such intent. In *DeCanas*, the Court noted that California Labor Code section 2805, which prevented an employer from knowingly employing an illegal alien if such employment would have an adverse effect on lawful resident workers, was constitutional as a regulation of immigration even though it dealt with noncitizens,[120] because states have broad authority under their police power to regulate employment relationships to protect workers within the state.[121]

Additionally, assuming that increased driver's license restrictions do come under the field of immigration, advocates of further restriction would argue either that the regulation was within their police power or that even if it were not within their police power, the increased restrictions do not burden the entrance or residence of noncitizens as prescribed.[122] If either case were proven, then advocates would have support that a complete ouster has not taken place.[123] Advocates would further contend that the restrictions do not burden the entrance or residence of noncitizens. In their support they might contend that not all noncitizens are prevented from obtaining a driver's license. In fact, since the number of lawfully present noncitizens who would be denied a license is likely to be such a small number, it would definitely not be a burden to entrance, and would not likely be a burden to residence.

Opponents of the increased restrictions would first counter by claiming that the increased restrictions do not come within the police powers of the states. *DeCanas* dealt with employment and illegal noncitizens, not the denial of a driver's license to lawfully present noncitizens.[124] In that case, it was clear that the employment of illegal noncitizens in a tight economic time could have an adverse economic impact on the citizens. With the current regulations, it is not clear what, if any, adverse economic impact would take place by allowing certain classes of noncitizens to obtain a driver's license.

Opponents might also argue that the regulation would "burden the entrance or residence of noncitizens."[125] States that have a "lawful presence" requirement clearly do burden immigrants and noncitizens. If noncitizens are unable to obtain a driver's license then they cannot freely commute to work, the store or any other

---

119. DeCanas v. Bica, 424 U.S. 351, 357 (1976) (quoting Rice v. Santa Fe Elevator Corp., 331 U.S. 218, 230 (1947)).

120. *Id.* at 365.

121. *See id.* at 356-57. Because illegal immigrants were the subject of the regulation, the Court did not even touch on whether the regulations burdened the entrance or residence of noncitizens.

122. *See DeCanas*, 424 U.S. at 356-57.

123. *Id.*

124. *Id.*

125. Michael S. James, *Parked: Rhode Island Foreign Students Unable to Obtain New Driver's Licenses*, ABC NEWS.COM, Feb. 11, 2002, *at* http://www.abcnews.go.com/sections/us/Daily/news/drivers_licenses020211.html.

place they may need to travel. Opponents argue that it should not matter how many noncitizens are affected, so long as there is a potential that the class may be burdened. They also point to the increase in the number of required documents necessary to obtain a driver's license and label this practice a state-imposed discriminatory burden.[126] It is not clear why requiring additional documents prevents those who are lawfully present in the United States according the federal government from obtaining a driver's license. If the federal government is satisfied that the person has proven his or her identity and reason for being in the country, why should states be permitted to deny them a driver's license? This is clearly a discriminatory burden.

In regards to the second way of showing preemption, a court would likely find that the current driver's license restrictions are permissible regulations of immigration. However, it is not clear whether there is a showing that it was the "clear and manifest purpose of Congress"[127] to effect a "complete ouster of state power, including state power to promulgate laws not in conflict with federal laws."[128] Advocates and opponents of the restrictions offer strong arguments for the legitimacy of their position. What will likely be the deciding factor under this analysis is the current attitude of the nation with respect to noncitizens. In light of the events September 11, a court would likely side with tougher restrictions even if the rights of some immigrants and noncitizens are trampled.

3. Congress May Preempt State Law That Conflicts with Federal Law.—The last way to demonstrate that a state law is preempted is if it "stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress."[129] Stated differently, a statute is preempted under the third scenario if it conflicts with federal law, making compliance with both state and federal law impossible.[130] If a court determines that driver's license restrictions are an enactment of immigration policy by the states, then the regulations would be preempted. However, it is not certain how a court would rule on this issue. Courts have dealt with the rights of illegal immigrants to driver's licenses,[131] but none has specifically dealt with the rights of lawfully present noncitizens and their access to driver's license.

It is possible that courts could follow the line of reasoning extended by the U.S. Supreme Court in *DeCanas*, finding validity of the state action under the police power.[132] If this view is accepted then the validity of the regulations would be upheld. However, if courts followed the line of reasoning extended by the

---

126. State regulation not congressionally sanctioned that discriminates against noncitizens lawfully admitted to the United States is impermissible if it imposes additional burdens not contemplated by Congress. 8 U.S.C. § 1101 (2003).

127. *DeCanas*, 424 U.S. at 357 (quoting Rice v. Sante Fe Elevator Corp., 331 U.S. 218, 230 (1947)).

128. *Id.*

129. *Id.* at 363 (citing Hines v. Davidowitz, 312 U.S. 52, 67 (1941)).

130. Mich. Canners & Freezers Ass'n v. Agric. Mktg. & Bargaining Bd., 467 U.S. 461, 469 (1984); *see also* Fla. Lime & Avocado Growers, Inc. v. Paul, 373 U.S. 132, 142-43 (1963).

131. Doe v. Ga. Dep't of Pub. Safety, 147 F. Supp. 2d 1369 (11th Cir. 2001).

132. *DeCanas*, 424 U.S. at 356-57.

U.S. Supreme Court in *Graham* that struck down state welfare laws which conditioned benefits on citizenship, then the validity of the regulations would be preempted.[133]

IV. POLICY REASONS WHY INCREASED RESTRICTIONS TO DRIVER'S LICENSES ARE AN INEFFECTIVE WAY TO PREVENT TERRORISM AND IDENTITY FRAUD

States' justifications for restrictions on driver's licenses are the prevention of terrorism and identity theft.[134] While increasing national security is critical, restricting driver's licenses is an ineffective way to prevent terrorism and identity theft. In fact, it is conceivable that denying driver's licenses to lawfully resident immigrants or noncitizens makes everyone less safe. Four guiding principles should be considered by states when drafting or modifying legislation pertaining to driver's license restrictions.[135]

A. Are the Restrictions Effective?

The first principle is that the driver's license restrictions must be effective.[136] Do the restrictions actually make us safer or are they just giving us a false sense of security? Is the proposal cost-effective or would a great amount of resources be expended for uncertain results? The ultimate question is whether the proposals achieve what was intended. The answer to this question is an emphatic no.

1. Ineffective in Their Prevention of Terrorism.—Restricting driver's licenses is an inefficient and ineffective measure to prevent terrorism. While it is true that many of the September 11 hijackers had obtained driver's licenses, the fact remains that they could have achieved their purposes without the licenses. Each hijacker had a foreign passport that allowed them to board the plane. It is unlikely that the federal government would not continue to allow passports as valid forms of identification to board planes because the United States receives a large volume of tourists and other visitors.

2. Ineffective Because of a Supremacy Clause Claim.—The license restrictions are also ineffective because of a supremacy clause claim, which renders the restrictions ineffective. There are three ways to show preemption in a supremacy clause analysis. The most likely means of preempting driver's license restrictions is the second test. Even if the state law is a permissible regulation of immigration, it may still be preempted if there is a showing that it

---

133. Graham v. Richardson, 403 U.S. 365, 378 (1971). The state statutes at issue in the instant cases impose auxiliary burdens upon the entrance or residence of noncitizens that suffer the distress, after entry, of economic dependency on public assistance. State laws that impose discriminatory burdens upon the entrance or residence of noncitizens lawfully within the United States conflict with this constitutionally derived federal power to regulate immigration and have accordingly been held invalid.

134. *See supra* notes 35-79 and accompanying text.

135. Waslin, *supra* note 43, at 5.

136. *Id.*

was the "clear and manifest purpose of Congress"[137] to effect a "complete ouster of state power, including state power to promulgate laws not in conflict with federal laws"[138] with respect to the subject matter which the statute attempts to regulate.[139] "State laws which impose discriminatory burdens upon the entrance or residence of aliens lawfully within the United States conflict with this constitutionally derived federal power to regulate immigration, and have accordingly been held invalid."[140]

Because courts have tended to disfavor restrictions that burden the entrance or residence of noncitizens, it is possible that state-imposed restrictions could be preempted by Congress' intent to fill the entire area of regulation of immigrants and noncitizens. One story of how the restrictions discourage the entrance or residence of noncitizens came from a British woman.

Carol Thornley lives off the southern coast of England. She had visited Florida thirty-eight times in the past five years and was planning on spending a large portion of her retirement there. She is now considering selling her retirement home because of the new restrictions on U.S. visas and Florida driver's licenses. The new restrictions have left her feeling unwanted and angry.[141] In the case of Thornley and other immigrants, immigration officials can restrict visas to as little as thirty days. When the visa expires, so does the immigrant's driver's license. The process for getting a new license requires so much additional paperwork that it could take up to a month to receive it in the mail.[142] This is but one story of persons lawfully in the United States but prevented from receiving a license.

Some of the other groups of immigrants who are here lawfully but could still be denied a license under the current restrictions are: refugees of special humanitarian concern, admitted into the United States;[143] persons who have been granted withholding of removal because of a likelihood of harm if returned to their country; asylees seeking refuge in the United States;[144] and professionals in a specialty occupation[145] who originally entered the United States as visitors and

---

137. *DeCanas*, 424 U.S. at 357 (quoting Rice v. Santa Fe Elevator Corp., 331 U.S. 218, 230 (1947)).

138. *Id.*

139. *Id.*

140. Takahashi v. Fish & Game Comm'n, 334 U.S. 410, 419 (1948); *see also* Hines v. Davidowitz, 312 U.S. 52, 66 (1941).

141. Melissa Harris, *Red Tape Frustrates Foreign Drivers; New Restrictions Since Sept. 11 Are So Inconvenient for Immigrants That Some May Drive Without a Valid State License*, ORLANDO SENTINEL, Aug. 19, 2002, at B1.

142. *Id.*

143. *See* 8 U.S.C. § 1157 (2003).

144. *See id.* § 1158(a).

145. *Id.* § 1101(a)(15)(H)(i)(b). This statute reads,

> having a residence in a foreign country which he has no intention of abandoning who is coming temporarily to the United States to perform other temporary service or labor if unemployed persons capable of performing such service or labor cannot be found in this country, but this clause shall not apply to graduates of medical schools coming to

then changed status. Persons in these groups will not have a visa, but are nonetheless lawfully present in the United States.

It is clear that driver's license restrictions can place a heavy burden on the entrance or residence of immigrants or noncitizens in the United States. However, even if a court rules that the restrictions are not preempted, the constitutionality of the restrictions should not encourage states to further support bad policy. If a court deals with unlawfully present immigrants under a preemption analysis, it is nearly certain that the court will not preempt the restrictions imposed by states.

B. Do the Restrictions Create Negative Unintended Consequences?

When states imposed the current restrictions aimed at preventing terrorism and identity theft, they did not do so with the intent of preventing legal immigrants from obtaining a license. Unfortunately, groups of these persons have been excluded. Sevilla Sanz, for example, was a twenty-eight year old international student who was unable to obtain a U.S. driver's license after her international driver's license expired. Although she passed the written and road tests, she was still denied. "They said basically if I was an international student I couldn't get my license."[146] Foreign students at the University of New Hampshire must travel (without a license) to the state capital to obtain a license, rather than to their local motor vehicle office. Once there, foreign students without valid foreign driver's licenses must obtain official forms from their home countries asserting they have never had a suspended license or revoked license there.[147] Sevilla Sanz is a representative of just one group of many affected by the restrictions. Many more excludable groups exist.

Driving is a necessity for many persons residing in the United States. Immigrants and noncitizens who are denied licenses still must find transportation to work, stores and school. Lawmakers would be naive to believe that all of these persons simply accept that they are not permitted to drive. Many of these persons drive unlicensed, ignorant of state driving rules and unlikely to be insured. The ultimate state interest in licensing drivers is "the important public interest in safety on the roads and highways, and in the prompt removal of a safety hazard."[148] The combination of unlicensed, uneducated and uninsured drovers not only significantly increases the number of traffic violations and accidents but also causes $4.1 billion in insurance losses per year.[149]

---

the United States to perform services as members of the medical profession; or (iii) having a residence in a foreign country which he has no intention of abandoning who is coming temporarily to the United States as a trainee, other than to receive graduate medical education or training; and the alien spouse and minor children of any such alien specified in this paragraph if accompanying him or following to join him.

Id.

146. James, *supra* note 125.

147. *Id.*

148. Dixon v. Love, 431 U.S. 105, 114 (1977).

149. *State Issues: Immigrants—Driver's Licenses*, STATE ACTION, *available at* http://www.

Lastly, by making people supply their SSNs, the risk of identity theft is increased, not decreased. Dishonest employees, like those who were caught in New Jersey,[150] can obtain the SSNs in the workplace and use them to apply for credit and assume the person's identity. States could allow individuals to apply for driver's licenses without having to provide a SSN; substitutions of an Individual Taxpayer Identification Number or an L-676 letter would suffice.[151]

It is self-evident that states would not intentionally prevent persons lawfully present in the United States from receiving a driver's license, increase hazards on the roadways, or make it easier for an identity thief to steal someone's SSN. Unfortunately, the states' choices did not have the intended effect but instead created negative consequences.

C. Do the Restrictions Single People out for Abuse and Discrimination?

To determine if people are being singled out for abuse and discrimination, a constitutional analysis will be applied. As noted in Part III of this note, an equal protection claim will be governed by the standard that "the governmental interest claimed to justify the discrimination is to be carefully examined in order to determine whether that interest is legitimate and substantial, and inquiry must be made whether the means adopted to achieve the goal are necessary and precisely drawn."[152] If departments of motor vehicles are attempting to achieve their policy justifications by insuring identity, it is unclear why lawfully resident noncitizens who possess documentation that satisfies federal government requirements, are potentially more of a hazard to license than a United States citizen. The rule is not narrowly tailored to meet the state's interest in assuring safety on the roads.

If states, by way of increased restrictions, are attempting to verify identification for the prevention of either terrorism or identity theft, they may be performing a function reserved for the federal government. "Control over immigration and naturalization is entrusted exclusively to the Federal Government, and a State has no power to interfere."[153] Once again, the statute is not narrowly tailored to meet the state's interest. However, it is conceivable that a court could find that restricting driver's licenses is a legitimate state interest and the means adopted to achieve the goal are necessary. If such a position was taken by the courts then the restrictions would be upheld.

D. Are the Regulations Based on Accurate Information?

Immigrant restrictions do not address the issue of false documents.[154] States' second justification for restrictions on driver's license is the need to prevent

---

cfpa.org/issues/immigrantdriving/index.cfm.

150. *See supra* note 66.

151. Waslin, *supra* note 43, at 12. An L-676 letter can be obtained by persons who can prove their age, identity and ineligibility to obtain a SSN.

152. Nyquist v. Mauclet, 432 U.S. 1, 7 (1977) (quoting Examining Bd. v. Flores de Otero, 426 U.S. 572, 605 (1976)).

153. *Id.* at 10.

154. Waslin, *supra* note 43, at 11.

identity theft by providing to appropriate law enforcement officials the ability to verify the authenticity of a license document, as well as identity of the license holder. The information on a driver's license is only as good as the information provided. Preventing immigrants and noncitizens from obtaining the licenses to which they are entitled encourages the use of false documents. This not only negates the policy justification but creates a false sense of security without addressing the real issues of identity fraud and theft.

E. Scope of the Solution

Safety and security goals are not mutually exclusive. Both can be accomplished through initiatives that combine effectiveness, accuracy, explicit civil rights protections, and prevention of discriminatory efforts.[155] "By arguing that driver's licenses are more than they are and that any possible 'fix' to the current driver's license system must address a range of unrelated collateral matters, advocates for an enhanced 'identity security' system cloud the real issues and hamper an appropriate and effective solution."[156] States should be striving to maintain the integrity of the license issuance and the verifiability of the license.

Better license management yields greater public safety and will improve the license use for other permissive uses. Steps must be taken to ensure that new policies are effective and make the country safer. Areas for state improvement include, but are not limited to, non-discriminatory issuance standards; driver information, including collection, sharing and exchange; state operation and enforcement; tamper- and counterfeit-proof features; accurate and reliable personal identifiers; and verifiability (enhancing communications and information infrastructure to allow real time access to driving history and authenticity of driver's license documents).[157]

CONCLUSION

After the September 11 attacks, the United States, collectively as a federal government and individually as states, became aware of potential security weaknesses. To rub salt in a devastating wound, it was discovered that several of the terrorists who wreaked havoc on this nation used our licensing system to help advance their purposes. This was a tragedy; and future attacks need to be stopped. On another front, invisible enemies, identity thieves, are attacking the citizens of this country. People's lives are being stolen and ruined by the simple act of SSN theft. These crimes, which are increasing in frequency, must be stopped.

The driver licensing system in the United States needs help. States believed they could alleviate the problems of terrorism and identity theft by increasing driver's license restrictions, but they were wrong. As a result of hasty action on the part of the states, Equal Protection Clause and Supremacy Clause violations

---

155. *Id.* at 14.

156. Harberson & Doherty, *supra* note 3, at 7.

157. *Id.* at 12.

have potentially risen. A state-issued driver's license offers proof of authorization to drive a motor vehicle in this country. Expanding the use of the license to prevent security concerns is not wise, especially when segments of the population, who are present lawfully in the United States, are prevented access to a license by such security concerns.

In enacting policies concerning driver's licenses, states should ask four questions: 1) are the restrictions effective?; 2) do the restrictions create negative unintended consequences?; 3) do the restrictions single people out for abuse and discrimination?; and 4) are the regulations based on accurate information? When states can answer no to all of these questions, the driver's license restrictions will be hurting terrorists, not immigrants and noncitizens.

THE UNTOUCHABLES: PROTECTIONS FROM LIABILITY FOR BORDER SEARCHES CONDUCTED BY U.S. CUSTOMS IN LIGHT OF THE PASSAGE OF THE GOOD FAITH DEFENSE IN 19 U.S.C. § 482(b)

NATHANIEL SAYLOR*

INTRODUCTION

Due to the increased public focus on security observable since the attacks of September 11, 2001, there is an ongoing debate concerning the proper balance to be struck between respecting civil rights and the role of the government in protecting the nation from terrorism. The growing concern about who and what is entering the country has provided proponents of stronger police powers with support in their efforts to increase that power. Nowhere is this debate more obvious than at the nation's airports and border crossings. The public is interested in seeing people and contraband linked to terrorism stopped before they enter the country. The bulk of the responsibility for protecting these public interests falls on the U.S. Customs Service ("Customs"). Customs officially transferred to the Department of Homeland Security on March 1, 2003.[1] The division of Customs that is charged with protecting the nation's borders and collecting duties on imported merchandise is now referred to as the Bureau of Customs and Border Protection.[2] Customs performs these duties by searching individuals and merchandise that enter the country. Of course, to carry out that mission Customs employees must make important decisions about the balance that must be struck between respecting civil rights and protecting the nation's borders. Customs has the authority to detain and search individuals and their property when entering the country, and because of the important public policies at issue, there is little doubt that Customs should have significant leeway in making those decisions. For this reason, Customs has been granted protections from liability by both Congress and the courts.

To understand the scope of these protections, an illustration is in order. Imagine that you are a female returning to the United States after a trip to Jamaica. After your flight arrives, a Customs narcotic detection dog "alerts" to

---

* J.D. Candidate, 2005, Indiana University School of Law—Indianapolis; M.B.A. Candidate, 2005, Kelley School of Business; B.M., 2001, Ball State University.

1. H.R. DOC. NO. 108-32, at 4 (2003).

2. According to the "Reorganization Plan Modification for the Department of Homeland Security," the U.S. Customs Service has been renamed to the "Bureau of Customs and Border Protection. *Id.* Further, the Homeland Security Act makes it clear that all references to any agency in a statute, regulation, directive, or delegation that was in place before the transfer will continue to apply to that agency, regardless of name changes, after the switch to the new department. Pub. L. No. 107-296, § 1512(d), 116 Stat. 2135 (2002). Therefore, all statutory protections afforded to employees of the former Customs Service will now apply to employees of the Bureau of Customs and Border Protection. Throughout this Note, all references to Customs will be understood as references to the Bureau of Customs and Border Patrol, which also includes former members of the Border Patrol.

your luggage. You are subsequently taken to a room where a pat down search is conducted by a female Customs inspector, revealing what the inspector interprets as a bulge in your pelvic region. The inspector then performs a partial strip search of your person. No drugs are found on you, or in your luggage, and you are released without further incident. On these same facts, a court recently, and not surprisingly, held that the inspector involved was not liable for damages for violation of the plaintiff's constitutional rights.[3] This result, which will be explained in more detail throughout this note, is not particularly surprising, or disturbing, because legitimate policies exist to justify the result. One such policy is to allow the government to protect its borders as the sovereign.[4] Without this protection, terrorist contraband and drugs would enter this country unchecked. However, when considering how quickly such an intimate search can cross the boundaries of acceptance, the question arises as to whether Customs should be granted more protections in order to avoid liability for possible digressions. The relevance of this question is especially appropriate in the aftermath of the events of September 11, 2001, and the ensuing "war on terror." Congress answered the question: more protection is exactly what Congress granted to Customs in August 2002.

In the summer of 2002, the Trade Act of 2002 became law.[5] Section 341 of the Act amends 19 U.S.C. § 482 by adding subsection (b). Before the amendment, 19 U.S.C. § 482 allowed any authorized person to board or search vessels to search for any merchandise which was subject to duty and/or being imported into the country contrary to law.[6] This section gave Customs inspectors the ability to conduct searches of "any vehicle, beast, or person"[7] at the nation's borders which they suspect has merchandise that is subject to duty, or that is being imported contrary to law.[8] The additional subsection grants persons conducting searches under § 482's authority immunity from liability in suits stemming from such searches if the person acts in good faith.[9] The provision is a radical departure from the traditional protections granted to Customs and was hotly debated in Congress. The amendment is the first instance in which Congress has enacted a statute that changes the standard for qualified immunity in a constitutional tort case.[10]

On its face, this amendment seems like it would be a radical shift from traditional ideas of liability and would go far in giving Customs inspectors a license to use and abuse their authority to search for contraband. Surprisingly, or perhaps disturbingly, because of the array of protections already available to Customs, the amendment will likely only be necessary in a limited number of

---

3. Saffell v. Crews, 183 F.3d 655, 659 (7th Cir. 1999).

4. The sovereign can protect itself by stopping and searching people and property entering the country. *Id.* at 657 (citing United States v. Ramsey, 431 U.S. 606, 616, 619 (1977)).

5. Pub. L. No. 107-210.

6. 19 U.S.C. § 482(a) (2003).

7. *Id.*

8. *Id.*

9. *Id.* § 482(b).

10. 148 CONG. REC. H5980 (daily ed. July 26, 2002) (statement of Rep. Conyers).

cases—those with egregious facts. As previously illustrated, Customs already enjoys numerous protections, with little concern of liability. This begs the question of what kind of conduct the passage of the recent amendment will end up protecting that was not protected under the previous framework. Indeed, the defense of qualified immunity, to be discussed later, already protects "all but the plainly incompetent or those who knowingly violate the law."[11]

Part I of this Note provides an overview of the protections already available to both Customs and its employees for searches prior to the enactment of the Trade Act of 2002. In Part II of this Note, I discuss good faith in general and the amendment to 19 U.S.C. § 482. The legislation itself gives no indication as to what will constitute good faith; therefore, I speculate on some factors that courts would be likely to consider as proof of good faith. Finally, I conclude by arguing that the amendment is overreaching, and because of the factual inquiries involved in establishing good faith, it will only be necessary when the pre-existing defense of qualified immunity fails, giving those incompetent and malicious individuals mentioned above one more opportunity to escape liability.

It is important to point out that this Note does not focus on searches of merchandise at sea ports and warehouses. The analysis centers around searches of individuals at airports and border crossings. Having said that, it is also important to note that 19 U.S.C. § 482 does give Customs the authority to search vessels and warehouses at sea ports. Therefore, all of the defenses to be discussed will be available to inspectors in suits arising from these latter types of searches. However, because of the more prevalent concern over individual civil rights that has arisen since September 11, 2001, this Note focuses on personal searches.

I. STATUTORY PROTECTIONS FROM LIABILITY

There exist a variety of avenues to seek redress from the government for harms caused by the government. The Federal Tort Claims Act (FTCA) allows suits against the government for some tortious conduct, while suits for violation of Fourth Amendment rights can be brought under *Bivens*.[12] As this section will illustrate, a recovery against the government under the FTCA is highly unlikely in a case involving Customs.

A. Federal Tort Claims Act

The FTCA represents a dramatic waiver of the sovereign immunity enjoyed by the government. Under the FTCA, the government is liable for the actions of its employees if the employee can be held liable under the law of the state where the claim arose.[13] One would assume that the FTCA would be a powerful

---

11. Courson v. McMillian, 939 F.2d 1479, 1487 (11th Cir. 1991) (quoting Malley v. Briggs, 475 U.S. 335, 341 (1986)).

12. Bivens v. Six Unknown Named Agents of Fed. Bureau of Narcotics, 403 U.S. 388 (1971) (holding that the performance of a search that does not comply with the requirements of the Fourth Amendment creates a cause of action for the aggrieved party).

13. 28 U.S.C. § 2672 (1994).

weapon for a person who was detained by Customs unjustly or whose property was damaged during a search by Customs. However, the FTCA has a number of exceptions, some of which apply to Customs and protect Customs from much of the liability that would normally be associated with these searches.

1. Discretionary Function Exception.—One exception to the FTCA that seems to be invoked more than any other is the discretionary function exception. This provision exempts the United States from FTCA liability for claims that are based upon the exercise, performance, or failure to exercise or perform any discretionary function or duty by a federal agency or employee, whether or not the discretion involved is abused.[14] The purpose of the discretionary function exception is to protect legislative and administrative decisions with social, economic, and political policy justifications from judicial second-guessing.[15] Not only are decisions establishing programs and implementing regulations protected, also protected are decisions by employees in exercising discretion allowed by those regulations.[16] This purpose seems to be directly related to Customs' decision of whether or not to search an individual, because that decision is an administrative decision backed by social policy (the policies of preventing entry of illegal contraband and protecting citizens from terrorist acts). Whether a court will find that Customs' decision to search an individual is protected by the discretionary function exception will differ depending on the jurisdiction in which the claim is adjudicated.[17] Wherever the claim is adjudicated, at least one fact seems clear, the discretionary function exception does not protect conduct that violates a legal mandate.[18] Therefore, a plaintiff can still recover for a search that is found to violate some established legal rule, even if a court finds that the Customs employees were performing a discretionary function.

Perhaps the most difficult part of applying the discretionary function exception is determining just what is a discretionary function. What factors will

---

14. *Id.* § 2680(a).

15. O'Ferrell v. United States, 253 F.3d 1257, 1266 (11th Cir. 2001) (citing United States v. Gaubert, 499 U.S. 315, 323 (1991)).

16. *Gaubert*, 499 U.S. at 323-24. "[I]f a regulation allows the employee discretion, the very existence of the regulation creates a strong presumption that a discretionary act authorized by the regulation involves consideration of the same policies which led to the promulgation of the regulations." *Id.* at 324.

17. *See, e.g.*, Nurse v. United States, 226 F.3d 996, 1002 (9th Cir. 2000) (stating that claims for false imprisonment, invasion of privacy, and negligence do not seem to be protected by the discretionary function exception); Caban v. United States, 671 F.2d 1230, 1233 (2d Cir. 1982) (holding detention of suspect at international airport did not "involve a choice between competing policy considerations"). *But see, e.g.*, Attallah v. United States, 955 F.2d 776, 784 (1st Cir. 1992) ("[A]ccording to the statute and regulations, the agents are not obligated to stop and search every passenger. This function, we believe, is a discretionary function as defined by § 2680 of Title 28 of the United States Code."); Jackson v. United States, 77 F. Supp. 2d 709, 714-15 (D. Md. 1999) (holding detention by Customs officials involved policy considerations and fell within discretionary function exception).

18. *Nurse*, 226 F.3d at 1002 (citing United States Fid. & Guar. Co. v. United States, 837 F.2d 116, 120 (3d Cir.), *cert. denied*, 487 U.S. 1235 (1988)).

a court consider to determine whether a discretionary function exists? As the name implies, for the discretionary function exception to apply, there must be room for the official to exercise discretion in making the decision at hand. The court in *Nurse v. United States*[19] set forth a test for determining when an act qualifies for protection as a discretionary function when it stated:

> In order to determine whether the discretionary function exception applies, the court must engage in a two-step inquiry. First, the court must determine whether the challenged conduct involves an element of judgment or choice.[20] Second, if the conduct involves some element of choice, the court must determine whether the conduct implements social, economic or political policy considerations.[21]

It would appear that the decision of a Customs inspector to select a certain individual for a search at an airport, or a certain vehicle at a port of entry, would qualify as a discretionary function under this test, and, as we have seen, some courts have so held. Therefore, for actions brought under the FTCA, we see that Customs may avoid liability if a court were to decide that the employees were performing a discretionary function when determining which passengers to search.

2. Detention Exception.—Regardless of whether the discretionary function exception will apply to a search, Customs has an even more powerful weapon available in the aptly named "detention exception." The detention exception, also known as the "law enforcement exception," provides protection for Customs officials from claims "arising in respect of" detention of goods by a Customs officer, as well as certain claims against "any other law enforcement officer."[22] This section has been interpreted to mean that Customs is not liable for any claim "arising out of" the detention of goods.[23] This interpretation has allowed courts to extend the detention exception beyond cases where Customs damages an individual's goods during an inspection. The exception has even allowed inspectors to avoid liability in cases where inspectors have caused physical harm to citizens so long as the harm occurred incident to a detention of property by Customs.[24] In *Jeanmarie*, the detention exception was stretched so far as to

---

19. 226 F.3d at 996.

20. *Id.* at 1101 (citing Berkovitz v. United States, 486 U.S. 531, 536 (1988)).

21. *Id.* (citing Gasho v. United States, 39 F.3d 1420, 1435 (9th Cir. 1994)).

22. 28 U.S.C. § 2680(c) (2000). This exception will continue to apply to Customs searches at the border, even though the name of Customs has changed, because of the Savings Provision of the Homeland Security Act which allows references to agencies in statutes to carry over despite the change in nomenclature. Pub. L. No. 107-296 § 1512(d), 116 Stat. 2135 (2002).

23. Kosak v. United States, 465 U.S. 848, 854 (1984). "But we think that the fairest interpretation of the crucial portion of the provision is the one that first springs to mind: 'any claim arising in respect of' the detention of goods means any claim 'arising out of' the detention of goods. . . ." *Id.*

24. Jeanmarie v. United States, 242 F.3d 600, 605 (5th Cir. 2001).

> [N]otwithstanding the fact that intentional tort claims arising out of arrests are not barred by § 2680(c), and are in fact permitted by § 2680(h), such claims *are barred* by

protect inspectors who physically restrained and injured a man who was looking for the restroom. The inspectors were searching his car and had told him to stay where he was. When the man left the inspection area anyway (because he had a medical condition making his use of the restroom urgent), the inspectors found him and physically restrained him, causing several injuries.[25] However, the court declined to hold the inspectors or Customs liable for the injuries, arguing that the detention exception barred recovery.[26] The detention exception is especially useful for Customs because almost all of Customs' interactions with the public occur as a result of a search of some type. For this reason, the detention exception will protect them from liability in the vast majority of suits brought under the FTCA.

The FTCA protects Customs from liability for the acts of its employees either when a court determines that the employees were exercising a discretionary function or when any claims arise during the detention of goods. As we have seen, the courts are willing to consider an inspection of a vehicle at the border as a detention,[27] and of course any search at an airport will be considered as arising out of the detention of goods because luggage is present and under Customs' control during personal searches. Therefore, it is difficult to envision a situation where a person being searched by Customs could argue that the detention exception did not apply. If they could, the court may decide to apply the discretionary function exception to avoid liability for Customs.

II. COMMON LAW PROTECTIONS FROM LIABILITY

Congress is not the only entity that has acknowledged the need for protections for government actors performing in their official capacities. The courts have also proved willing to provide protections for these same actors. Customs inspectors, like most government agents, are entitled to "qualified immunity" for actions taken in their official capacities. Qualified immunity is a powerful defense for government employees, discharging liability for all but the most heinous of actions. Even though qualified immunity allows some seemingly heinous conduct, there are important policy justifications for its continued existence. Following is a brief synopsis of the development of qualified immunity and its general application. A general understanding of qualified immunity is important before the scope of the "border authority"

> the customs-duty exception if the alleged torts arose from the inspection, seizure, or detention of goods by a Customs agent because such claims involve conduct covered by § 2680(c).

Id. at 604 (citing *Gasho*, 39 F.3d at 1433-34).

25. *Id.* at 601-02.

26. *Id.* at 604. Interestingly, the *Jeanmarie* court held that 28 U.S.C. § 2680(h), which does allow certain claims arising from intentional torts by other law enforcement officers notwithstanding the application of 28 U.S.C. § 2680(c), does not allow claims for those same intentional torts if the officer involved is a Customs employee to whom the detention exception would normally apply. *Id.*

27. *Id.*

protection enjoyed by Customs, discussed later, can be fully appreciated.

A. Qualified Immunity

Qualified immunity has developed as a defense to protect government actors from liability for suits brought against them for actions taken in their official capacities. The doctrine is not available to government employees who knew or should have known that the actions they took violated the constitutional rights of the plaintiff, or if the act was performed with malicious intention to deprive constitutional rights or cause some other injury.[28] Obviously, qualified immunity has a subjective aspect. The Supreme Court summed up the standard in 1974: "It is the existence of reasonable grounds for the belief formed at the time and in light of all the circumstances, coupled with good-faith belief, that affords a basis for qualified immunity of executive officers for acts performed in the course of official conduct."[29]

The Supreme Court altered its reliance on subjective good faith intention in 1982 in *Harlow*.[30] In that case, the Court listed several costs to society caused by extended litigation over subjective good faith including "distraction of officials from their governmental duties, inhibition of discretionary action, and deterrence of able people from public service."[31] Another important impetus for this change was to avoid waste of valuable government resources to defend mere allegations.[32] The Court justified the change when it stated that by relying on the objective reasonableness of an employee's actions, which would be measured against clearly established law, the courts could avoid excessive disruption of official duties, and allow many insubstantial claims to be resolved in summary judgment motions.[33] The Court also pointed out that "the public interest in deterrence of unlawful conduct and compensation of victims remains protected by a test that focuses on the objective legal reasonableness of an official's acts."[34] It is this test that courts apply to acts of Customs inspectors for claims arising against them in the scope of their employment, whether the inspector violated a clearly established legal mandate at the time of the actions. However, the question remained as to what was a clearly established legal mandate. The

---

28. Wood v. Strickland, 420 U.S. 308, 322 (1975).

29. Scheuer v. Rhodes, 416 U.S. 232, 247-48 (1974).

30. Harlow v. Fitzgerald, 457 U.S. 800 (1982).

31. *Id.* at 816.

32.

> [W]e conclude today that bare allegations of malice should not suffice to subject government officials either to the costs of trial or to the burdens of broad-reaching discovery. We therefore hold that government officials performing discretionary functions, generally are shielded from liability for civil damages insofar as their conduct does not violate clearly established statutory or constitutional rights of which a reasonable person would have known.

Id. at 817-18.

33. *Id.* at 818.

34. *Id.* at 819.

Supreme Court shed light on this question five years later in *Anderson v. Creighton*.[35]

The issue in *Anderson* was whether a federal agent who conducted a search that violated the Fourth Amendment could be held liable for money damages if a reasonable officer could have believed that the search did not violate the Fourth Amendment.[36] The Court noted that the *Harlow* standard depends on how generally the legal rule allegedly violated had been defined.[37] The Court realized that allowing the plaintiff to state the legal rule too broadly would annul the qualified immunity defense. A complaining party could simply allege that the defendant had violated a constitutional principle, and since constitutional principles, such as the protection against unreasonable searches, are largely established, the claim would have to go to trial and the resources that the Court enunciated in *Harlow* would be wasted. In addition to wasting resources by allowing claims based on very general principals to proceed to trial, an overly broad interpretation of the relevant legal principles also affects an official's ability to perform his or her job effectively.[38] In the end, the *Anderson* Court determined that the right must be so clearly defined that an official could tell that her conduct was illegal.[39]

Oftentimes, qualified immunity is referred to as qualified "good faith" immunity, but that term is misleading since there is no requirement of subjective good faith.[40] The *Anderson* Court did state that when determining whether the

---

35. 483 U.S. 635 (1987).

36. *Id.* at 636-37.

37. The Court gave the example:

> [T]he right to due process of law is quite clearly established by the Due Process Clause, and thus there is a sense in which any action that violates that [c]lause (no matter how unclear it may be that the particular action is a violation) violates a clearly established right. Much the same could be said of any other constitutional or statutory violation. But if the test of "clearly established law" were to be applied at this level of generality, it would bear no relationship to the "objective legal reasonableness" that is the touchstone of *Harlow*.

Id. at 639.

38. Allowing such a general definition of the relevant legal principle would destroy the balance that had been struck between the interest of citizens in vindicating constitutional rights and in officials performing their official duties because it would make it impossible for officials to determine when their conduct might give rise to liability. *Id.* at 639-40 (citing Davis v. Scherer, 468 U.S. 183, 195 (1984)).

39.

> The contours of the right must be sufficiently clear that a reasonable official would understand that what he is doing violates that right. This is not to say that an official action is protected by qualified immunity unless the very action in question has previously been held unlawful . . . but it is to say that in light of pre-existing law the unlawfulness must be apparent.

Id. at 640 (citing Mitchell v. Forsyth, 472 U.S. 511, 535 n.12 (1985)) (citing Malley v. Briggs, 475 U.S. 335, 344-45 (1986)) (citing Davis v. Scherer, 468 U.S. 183, 191, 195 (1984)).

40. Suissa v. Fulton County, Ga., 74 F.3d 266, 269 (11th Cir. 1996) ("Although the cases

actions of the official were objectively reasonable the information that was in the possession of the official at the time those actions were taken should be considered. However, this inquiry was not the same as the good faith inquiry that *Harlow* ended.[41] Courts continue to advance and alter the rule stated in *Anderson* to determine how to define the legal mandate at issue.[42] How a court defines the question of law at issue is very important to the disposition of a case where qualified immunity is at issue because of the procedural aspects of such a claim.

In order to fully understand the usefulness of the qualified immunity defense, one must understand its procedural context. As stated previously, many of the justifications for the defense of qualified immunity are to protect valuable government resources from time spent defending frivolous lawsuits. For this reason, qualified immunity is not just immunity from liability, but also immunity from suit.[43] Because the principal question in a qualified immunity case is whether the right allegedly violated was clearly established, and due to the fact that after *Harlow* there was no subjective element to the inquiry, these claims are often very amenable to disposition by summary judgment. This both protects government officials from extended litigation and saves government resources in defending such claims. In the end, whether the plaintiff is victorious will depend on how the court defines the legal question at issue because the court will then ask if that question has been definitively settled by existing case law.[44]

As illustrated, there are important policy justifications for the qualified immunity defense. Because of these justifications, the defense is a very powerful tool that government employees can use to not only avoid liability for acts committed in their individual capacities, but also to avoid suits by dismissing meritless claims at early stages of the litigation. Now that we have a basic understanding of the doctrine of qualified immunity as it applies to most government officials, we can consider certain other protections that make this defense even more useful for Customs employees.

B. Border Authority

Even before the passage of the Trade Act of 2002, which granted Customs inspectors immunity for searches conducted in good faith, Customs inspectors enjoyed considerably more freedom in performing searches than most other law

---

sometimes refer to the doctrine of qualified 'good faith' immunity, the test is one of objective legal reasonableness, without regard to whether the government official involved acted with subjective good faith.") (quoting Swint v. City of Wadley, Ala., 51 F.3d 988, 995 (11th Cir. 1995)).

41. *Anderson*, 483 U.S. at 641 (citing *Harlow*, 457 U.S. at 815-20). (The court concluded that "The relevant question in this case . . . is the objective . . . question whether a reasonable officer could have believed Anderson's warrantless search to be lawful. . . . Anderson's subjective beliefs about the search are irrelevant.") *Id.*

42. "If case law, in factual terms, has not staked out a bright line, qualified immunity almost always protects the defendant." Lassiter v. Ala. A&M Univ., Bd. of Trs. 28 F.3d 1146, 1150 (11th Cir. 1994) (quoting Post v. City of Fort Lauderdale, 7 F.3d 1552, 1557 (11th Cir. 1993)).

43. *Mitchell*, 472 U.S. at 526.

44. *Id.* at 535.

enforcement officers.[45] The reason for this freedom is past court decisions holding that the interests of the sovereign in protecting itself outweigh the privacy interests of an individual at the nation's borders.[46] This extended police power is commonly referred to as the "border authority." What follows is a summary of current border authority precedent. The border authority does not modify the qualified immunity inquiry. Instead, the border authority alters the prime facie showing of a constitutional violation, which is a pre-requisite to the court considering whether qualified immunity exists. For instance, the rules governing when a search will be considered unreasonable under the Fourth Amendment are different for claims arising on the national border than they are for claims arising on the interior. For this section, it is important to realize that international airports are considered national borders.[47] Therefore, the protections of the border authority are extended to Customs' work at these airports. Again, the border authority does not alter the qualified immunity analysis, but instead may affect when that analysis will be undertaken.

Many of the constitutional claims against Customs inspectors arise under the Fourth Amendment, which protects against unreasonable searches.[48] To understand how the border authority affects a finding of unreasonableness under the Fourth Amendment, a quick overview of the standard under a "normal" search is in order. To determine whether a "normal" search (i.e., a search that is not conducted at the border) violates the Fourth Amendment, the court will consider the scope of the intrusion, the way it was conducted, what the justifications are, and the place where it was conducted.[49] In order to decide whether a violation of the Fourth Amendment exists, the court weighs public interest against the Fourth Amendment interests of the individual.[50] The individual's interest in preserving privacy is well respected by the courts when the search was conducted in the nation's interior.

In contrast, courts have placed a much higher value on the sovereign's right to protect itself at the borders than on the individual's right to privacy. The executive has authority to conduct routine searches and seizures at the borders without probable cause or a warrant in order to collect duties and prevent the introduction of contraband.[51] The courts have determined that to accomplish this task some of the protections that citizens take for granted on the interior have to

---

45. United States v. Cascante-Bernitta, 711 F.2d 36, 37-38 (5th Cir. 1983).

46. United States v. Montoya de Hernandez, 473 U.S. 531, 539-41 (1985).

47. Almeida-Sanchez v. United States, 413 U.S. 266, 273 (1973).

48. U.S. CONST. amend. IV.

49. Bell v. Wolfish, 441 U.S. 520, 559 (1970). *See also* New Jersey v. TLO, 469 U.S. 325, 337-42 (1985) (holding what is reasonable depends on all the circumstances and the nature of the search and seizure).

50. United States v. Brignoni-Ponce, 422 U.S. 873, 878 (1975). *See also* Terry v. Ohio, 392 U.S. 1, 20-21 (1968).

51. United States v. Ramsey, 431 U.S. 606, 616-17 (1977). This authority comes from the Constitution Article I, Section 8, Clause 3, which grants power to prevent smuggling and to prevent prohibited articles from entry. *Id.* at 618-19 (citing United States v. 12 200-Ft. Reels of Film, 413 U.S. 123, 125 (1973)).

be lessened. Specifically, it has been held that routine searches at the border can be conducted without any requirement of probable cause.[52] Of course, there are valid reasons for this result, and upon even a little reflection one realizes that if there was a requirement of warrants at the borders, either trade would grind to a halt while investigations were conducted, or, more likely, contraband would simply enter the country unhindered. Obviously either result is unacceptable; therefore, the courts have placed the balance of interests in the government's favor as opposed to the individual citizen's right to privacy.

As stated, the *Montoya de Hernandez* Court determined that routine border searches did not require probable cause. The obvious next question is, what border searches are to be considered routine? The cases describe numerous types of searches performed by Customs, from luggage inspections and pat downs to strip searches and x-rays. Of course, not all of these searches can be considered routine. Therefore, not all are exempt from the probable cause requirement. Pat down searches seem to be the most prevalent types of searches that are litigated. This result is not surprising since pat down searches are a quick and easy way to search for merchandise. Of the circuit courts that have considered the issue, none have held that a pat down search is a non-routine border search.[53]

While legitimate reasons do exist for not requiring probable cause for routine searches at the border, the courts have granted some powers to inspectors that do not seem to have as much justification. For instance, the courts have even given Border Patrol agents the right to racially profile when determining which vehicles to stop at immigration checkpoints near the border.[54] These troubling holdings have been justified by the argument that race is a relevant factor in the search for illegal aliens.[55] While there is obviously a legitimate argument that race is an important factor in determining whether a person is an illegal alien,

---

52. United States v. Montoya de Hernandez, 473 U.S. 531 (1985). "[T]he Fourth Amendment's balance of reasonableness is qualitatively different at the international border than in the interior. Routine searches of the persons and effects of entrants are not subject to any requirement of reasonable suspicion, probable cause, or warrant." *Id.* at 538.

53. Bradley v. United States, 299 F.3d 197, 203 (3d Cir. 2002) (citing United States v. Bearas, 183 F.3d 22, 26 (1st Cir. 1999)) (citing United States v. Gonzalez-Rincon, 36 F.3d 856, 864 (9th Cir. 1994)) (stating luggage and pat down searches are routine and there is no requirement of reasonable suspicion to conduct them).

54. *Montoya de Hernandez*, 473 U.S. at 538. "Automotive travelers may be stopped at fixed checkpoints near the border without individualized suspicion even if the stop is largely based on ethnicity." *Id.* (quoting United States v. Martinez-Fuerte, 428 U.S. 543, 562-63 (1976)). *See also Martinez-Fuerte*, 428 U.S. at 563-64 (stating "even if it be assumed that such referrals are made largely on the basis of apparent Mexican ancestry, we perceive no constitutional violation") (citing *Brignoni-Ponce*, 422 U.S. at 878).

55. "To the extent that the Border Patrol relies on apparent Mexican ancestry at this checkpoint . . . that reliance clearly is relevant to the law enforcement need to be served." *Martinez-Fuerte*, 428 U.S. at 564 n.17. Justification for this statement came from the *Brignoni-Ponce* Court when it explained that the likelihood of any given person of Mexican ancestry being an alien was high enough to make race a relevant factor in determining which vehicles to stop for inspection. *Brignoni-Ponce*, 422 U.S. at 886-87.

there is also a point to be made that a more compelling public interest should be required before race is allowed as a criteria for conducting searches that lack probable cause. Simply trying to stop illegal aliens from entering the country may not seem like an issue that justifies such a blatant intrusion on an individual's rights.

Fortunately, and regardless of how one feels about the issue, it does not seem that Customs would be able to build on these prior holdings to make an argument that racial profiling in any other context is valid. The *Brignoni-Ponce* Court specifically stated that Mexican ancestry was a relevant factor for the Border Patrol to consider in determining whether a vehicle was carrying illegal aliens.[56] Customs would be hard pressed to argue similarly that race is a relevant factor in determining who is or is not attempting to bring dutiable goods or contraband into the country illegally.

One interesting side-note is that while these holdings dealing with immigration checkpoints are based on the border authority, the checkpoints at issue were not actually at the border. Therefore, Customs inspectors may enjoy the protections of the border authority beyond the actual border. This raises the question of where Customs authority to conduct searches without probable cause should cease, but that is a topic for another note.

One issue that is relevant to this note is the question left open by the *Montoya de Hernandez* Court, namely what level of suspicion, if any, is required for a non-routine border search. The Court listed strip searches, body cavity searches, and involuntary x-ray searches as non routine searches.[57] Of the courts that have considered what level of suspicion is required for these more intrusive searches, they have consistently held that reasonable suspicion is required.[58] Reasonable suspicion has been defined as "a particularized and objective basis for suspecting the particular person" of smuggling contraband,[59] and requires that the inspectors have a "particularized and objective basis for suspecting the particular person of . . . smuggling."[60] Moreover, courts have observed that when a search progresses from a routine border search to a non-routine search, the inspector must be able to justify the progression with reasonable cause as defined above.[61] Therefore, there is no requirement of suspicion for a routine search at

---

56. *Bringoni-Ponce*, 422 U.S. at 886-87.

57. *Montoya de Hernandez*, 473 U.S. at 541 n.4.

58. *Bradley*, 299 F.3d at 203 (citing *Gonzalez-Rincon*, 36 F.3d at 864) (citing United States v. Yakubu, 936 F.2d 939 (7th Cir. 1991)) (citing United States v. Oyekan, 786 F.2d 832, 837-39 (8th Cir. 1986)) (citing United States v. Carreon, 872 F.2d 1436, 1442 (10th Cir. 1989)).

59. United States v. Cortez, 449 U.S. 411, 417 (1981).

60. *Montoya de Hernandez*, 473 U.S. at 541-42 (quoting *Cortez*, 449 U.S. at 417 (1981)). *See also* Brent v. Ashley, 247 F.3d 1294, 1300 (11th Cir. 2001) (stating "[r]easonable suspicion to justify a strip search can *only* be met by a showing of articulable facts which are particularized as to the place to be searched") (citing United States v. Vega-Barvo, 729 F.2d 1341, 1349 (11th Cir. 1984)) (emphasis added).

61. "[A]s a search progresses from a stop, to a pat-down search, to a strip search, an agent must reevaluate whether reasonable suspicion to justify the next level of intrusion exists in light of the information gained during the encounter." *Ashley*, 247 F.3d at 1300 (citing *Vega-Barvo*, 729

the border, but non-routine searches require reasonable suspicion.

It is easy to lose sight of constitutional protections afforded individuals in the context of border searches. As stated earlier, the Fourth Amendment is designed to protect against unreasonable searches and seizures. How can a search conducted without any reasonable suspicion or probable cause fail to run afoul of the Fourth Amendment's mandate? The answer is that people at the border still have the right to be free from unreasonable searches and seizures, as guaranteed by the Fourth Amendment, but their expectation of privacy is less at the border than in the interior.[62] As stated earlier, the courts have struck the balance between the individual's right to privacy at the borders and the right of the government to protect itself in the government's favor.[63] In addition to the sovereign's interest in protecting itself, courts have pointed to the practical problems with a requirement of probable cause for searches conducted on inland routes into the country.[64] It is reasonable to assume that this same concern justifies the lack of a probable cause requirement for searches conducted at an airport as well.

The border authority is obviously an important doctrine. It is important not only for the protections that it affords inspectors from liability, therefore protecting valuable government resources from constant litigation, but also because it allows government actors to perform the important task of protecting the nation's borders without fear of personal liability. Of course, any discussion of the border authority involves important competing policies, such as the proper balance between an individual's civil rights and the government's ability to protect its borders. Now that these policies and arguments have been set out, it is possible to consider how these standards operate while taking a broad view of the policies driving their implementation.

III. THE STANDARDS IN ACTION

At this point, the majority of the defenses available to Customs employees, as well as the policies and arguments justifying those defenses, have been discussed. What remains, before a discussion of how those standards have been changed with the amendment to 19 U.S.C. § 482 should be undertaken, is a brief illustration of how those standards already operate in "real world" settings. To accomplish this illustration, this section will simply set out facts from recent cases, or facts that seem characteristic of a large percentage of cases in this area, and give an overview of how the discussed protections affect Customs employees.

In *Bradley*, the plaintiff alleged that the search Customs performed on her violated the Fourth Amendment.[65] Bradley had been subjected to a pat down.

F.2d at 1349).

62. *Montoya de Hernandez*, 473 U.S. at 539 (citing Carroll v. United States, 267 U.S. 132, 154 (1925)).

63. *Id.* at 540.

64. United States v. Martinez-Fuerte, 428 U.S. 542, 557 (1976).

65. Bradley v. United States, 299 F.3d 197, 203 (3d Cir. 2002).

She alleged that the search was an "intrusive patdown,"[66] and required reasonable suspicion. The court did not answer the question of whether an intrusive patdown would require some level of suspicion because it concluded that the patdown at issue was not intrusive and was still a routine border search.[67] The facts of the case were that Bradley was returning from Jamaica, a country which Customs considers a "source country" on a flight that Customs considers a source flight. The search involved a pat down over Bradley's dress. Bradley claimed that the inspector had inappropriately pushed on her breasts and inner and outer labia.

In its analysis of the allegations the court noted that searches at the border are presumed to be reasonable under the Fourth Amendment,[68] and that immigration checkpoints at international airports are the functional equivalent of national borders. The court went on to cite *Montoya de Hernandez* for the proposition that the sovereign has authority to conduct routine border searches without probable cause or warrant.[69] As such, the search was not subject to any requirement of suspicion. Therefore, Bradley's rights were not violated.[70] The court did not reach the qualified immunity issue because after analyzing the search under existing border authority precedent, the court determined that there was no constitutional violation.[71] It was further stated in dicta that even if there had been a constitutional violation, the inspectors would have been entitled to qualified immunity because there was no "clearly established" law holding that an intrusive pat down was not a routine search requiring some level of suspicion.[72] The court did not hold that qualified immunity protected the inspectors, but that under the border authority the search was permissible and therefore a qualified immunity analysis was not necessary.[73]

Now that we have seen an example of how all the protections available to Customs operate to protect them from liability, let us consider an example of a situation where the conduct was so egregious that the inspectors were not entitled to qualified immunity in the early stages of litigation. In *Brent v. Ashley*,[74] two Customs inspectors were challenging the denial of the motion for summary judgment, which was based on a theory of qualified immunity. According to the facts of the case, which consist of Brent's version, Brent was returning from Nigeria and on the final leg of the flight, she met another passenger named

---

66. *Id.* at 201.

67. *Id.* at 203.

68. *Id.* at 201 (citing United States v. Ramsey, 431 U.S. 606, 616 (1977)) (citing United States v. Hyde, 37 F.3d 116, 118-20 (3d Cir. 1994)) (citing United States v. Ezeiruaku, 936 F.2d 136, 140 (3d Cir. 1991)).

69. *Id.* at 201-02 (citing United States v. Montoya de Hernandez, 473 U.S. 531, 537 (1985)).

70. *Id.* at 204 ("We need not decide whether the customs inspectors reasonably suspected that Bradley was smuggling contraband because we conclude that the patdown was not so intrusive as to be transformed into a nonroutine border search.").

71. *Id.* at 205.

72. *Id.*

73. *Id.*

74. 247 F.3d 1294 (11th Cir. 2001).

Elbute. Brent and Elbute were the only black passengers on the flight. When they arrived, Customs searched Elbute and Brent displayed her disapproval. Based on her gesture, Inspector Seymour Schor instructed Inspector Carl Pietri to detain Brent who then alleged that she was being singled out because she was black.[75] After a detailed search of Brent's luggage, which returned nothing, the inspectors decided to detain Brent for further questioning. After questioning, the inspectors decided to conduct a full body pat down and a strip search. Three female inspectors were enlisted to conduct these searches. The form filed at the time listed the justifications for the search as merely nervousness and the fact that Brent had arrived from a "source country."[76] After the strip search, which included touching Brent's crotch area, removing her sanitary napkin, squeezing her abdomen, and monitoring her responses, Brent asked if she could use the restroom. Brent was allowed to use the rest room but was watched closely by inspectors and told not to flush the toilet so the inspectors could look for signs of contraband in her urine.

In spite of the fact that all the searches up to this point were negative, the original inspectors decided that an x-ray and pelvic exam should be performed at the hospital. Again, the justifications were nervousness and arrival from a source country. Brent was hand cuffed and presented with a consent form at the hospital, which she was told to sign or she would be held indefinitely. Again the searches were negative. Finally, ten hours after first being detained, Brent was released and allowed to return to Houston.[77]

On the above listed facts the district court granted the summary judgment motions of the subordinate inspectors, including the females who had conducted the strip search, but the motions of the two original inspectors were denied.[78] On appeal, the circuit court first addressed the question of whether Brent's constitutional rights had been violated. The court held that the initial stop was justified under the border authority, but that to hold a traveler beyond the scope of a routine search is only justified if, considering all the facts, the traveler was reasonably suspected of smuggling contraband.[79] Also, the court stated, "reasonable suspicion to justify a strip search can *only* be met by a showing of articulable facts which are particularized as to the place to be searched."[80] Further, as a pat down progresses to a strip search, the agent must reevaluate whether reasonable suspicion to justify the heightened intrusion exists, considering the information gained during the previous search.[81] The court, after considering similar cases with similar justifications, found that the only cause for suspicion of Brent was her nervousness, and that nervousness alone could not justify a strip search.[82] The court held that because the initial stop of Brent and

---

75. *Id.*
76. *Id.* at 1298.
77. *Id.*
78. *Id.*
79. *Id.* at 1299 (citing United States v. Montoya de Hernandez, 473 U.S. 531, 541 (1985)).
80. *Id.* at 1300 (citing United States v. Vega-Barvo, 729 F.2d 1341, 1345 (11th Cir. 1984)).
81. *Id.* (citing *Vega-Barvo*, 729 F.2d at 1349).
82. *Id.* at 1302 (citing United States v. Tapia, 912 F.2d 1367, 1371 (11th Cir. 1990)).

search of her luggage failed to produce a particularized or objective basis that would equate to particularized evidence that she was a drug courier, the search violated the Fourth Amendment.[83] For the same reasons, the court found that the x-ray was unconstitutional.[84] The court went on to hold that the supervisors were not entitled to qualified immunity because "a reasonable customs agent at the time of the incident would have known that a strip search under the facts of this case was a violation of Brent's Fourth Amendment rights."[85] As stated, *Brent* came to the court as an appeal of the district court's denial of a summary judgment motion. Therefore, even though the appellate court held the inspectors were not entitled to qualified immunity on the facts presented, which were the facts as stated by the non-moving party, the inspectors still had the opportunity to a trial in which they could discredit those facts.

To the reasonable individual, the conduct of the inspectors in *Brent* likely seems outrageous. The fact that the government even filed a motion for summary judgment on the facts of this case seems disturbing because it can be understood as evidencing a mentality that inspectors should not be held liable for their conduct. The obvious question then arises: why do Customs inspectors need more protection from liability, especially when that protection is seemingly premised on the subjective knowledge and intent of the inspector? Some will undoubtedly argue that Customs needs protection so they can conduct proper searches at the nation's borders without the fear of prosecution. While this is a valid argument, it does not explain why the protections already in place are insufficient. However, the debate about whether Customs should have more protection became largely moot in the summer of 2002 when Congress passed the amendment to 19 U.S.C. § 482. As stated previously, Customs now has immunity from liability for all searches conducted in good faith. Due to the recent passage of the act, however, there is no clear standard of what will suffice as good faith. The purpose of the following section is to offer some insight as to what the courts will likely consider as evidence of good faith.

IV. THE STANDARD FOR GOOD FAITH

No case law currently exists defining what will constitute good faith when the defense is raised by Customs for the first time. Further, because of the sheer size of the Trade Act of 2002, legislative history on the point is sparse at best. Therefore, to determine what the courts will consider as guidance, it is best to consider other statutory schemes with good faith defenses. However, the first relevant question is whether a good faith defense alters the effect of any of the existing defenses available to inspectors. In other words, will plaintiffs be able to argue successfully that by granting Customs inspectors good faith immunity from liability, Congress intended to deny Customs inspectors the existing defense of qualified immunity?

---

83. *Id.*

84. *Id.* at 1303.

85. *Id.* at 1305.

A. Effect of Good Faith Defense on Qualified Immunity

Whether good faith abrogates the common law defense of qualified immunity is a question that has recently been answered in the contexts of both the Fair Housing Act,[86] and the Federal Wiretap Act.[87] In *Gonzalez*, the court concluded that a good faith defense did not deny the defendant the protection of qualified immunity reasoning that "[n]either the text nor the legislative history of section 3617 [of the Fair Housing Act] indicates that Congress intended to abrogate the qualified immunity to which executive-branch officials were entitled under common law."[88]

The *Tapley* court went into more detail in its analysis of the good faith defense in the Federal Wiretap Act. The *Tapley* court rejected the notion that Congress had intended to abrogate the qualified immunity defense with the inclusion of a good faith defense when it stated, "the qualified immunity defense is so well-rooted in our jurisprudence that only a specific and unequivocal statement of Congress can abolish the defense."[89] The court also noted that the test for qualified immunity is objective, while the test for good faith is subjective.[90] It was also observed that one of the main benefits of qualified immunity, the fact that it helps to resolve claims early in the proceedings, does not exist in a good faith defense.

> We would not strip a judge or prosecutor of absolute immunity because the claim related to a statutory violation and the statute provided an affirmative defense. By the same token, police officers and public officials performing governmental functions should not lose their qualified immunity because of an affirmative defense which might or might not protect them but would, in all events, require they be subject to extended litigation and deprive them of the benefits of qualified immunity.[91]

This statement raises an interesting issue. Because a good faith defense rests on the subjective state of mind of the inspector, determining whether good faith exists will be a fact sensitive inquiry. As has already been illustrated, qualified immunity is often raised early in litigation before decisions of fact are resolved. Therefore, the good faith defense will only be required in cases where qualified immunity has failed, or when the qualified immunity question cannot be answered on summary judgment. As already illustrated in the previous section, there are very few cases where the qualified immunity defense does not protect inspectors, and cases where the inspectors are not entitled to qualified immunity seem to share fairly egregious facts. Therefore, the only effect of the good faith

---

86. Gonzalez v. Lee County Hous. Auth., 161 F.3d 1290 (11th Cir. 1998).

87. Tapley v. Collins, 211 F.3d 1210 (11th Cir. 2000).

88. *Gonzalez*, 161 F.3d at 1299.

89. *Tapley*, 211 F.3d at 1216 (citing Buckley v. Fitzsimmons, 509 U.S. 259, 268 (1993)).

90. *Id.* at 1215 (citing Anderson v. Creighton, 483 U.S. 635, 639 (1987)) (citing Harlow v. Fitzgerald, 457 U.S. 800, 816 (1982)).

91. *Id.* at 1216 (citing Blake v. Wright, 179 F.3d 1003, 1012 (6th Cir. 1999)).

defense will be to provide another layer of protection to those inspectors who were not protected by qualified immunity.

B. Defining Good Faith

As it appears that no court will find that the inclusion of a good faith defense abrogates qualified immunity, the next relevant question is how courts will define good faith. Initially, the standard of review in a good faith case has been stated as "whether a reasonable jury on the evidence adduced by the plaintiff and drawing all inferences in plaintiff's favor, could reasonably have found that the defendants acted in other than subjective good faith."[92] To determine what good faith is, it will be necessary to draw parallels from existing statutory schemes. Title III of the Omnibus Crime Control and Safe Streets Act of 1968, also known as the Federal Wiretap Act, contains a good faith defense.[93] This act allows someone who has been subjected to illegal surveillance to bring an action for damages, however, "[a] good faith reliance on . . . a court warrant or order, a grand jury subpoena, a legislative authorization, or a statutory authorization . . . is a complete defense against any civil or criminal action brought under this chapter or any other law."[94] For purposes of Title III, good faith consists of a subjective good faith belief that the official was acting in compliance with the statute, and that the belief was reasonable in and of itself.[95]

The *Kilgore* court went on to point out that "[i]f the requisites of the statutory good faith defense are met, then the standard for qualified immunity as a defense to Fourth Amendment violations is also satisfied."[96] However, this statement was made before the Supreme Court did away with the subjective good faith aspect of qualified immunity in *Harlow*.[97] Thus, it is not clear from this statement whether the court meant that a finding of statutory good faith satisfied the subjective aspect of qualified immunity in place at that time, or whether the court meant that upon a finding of statutory good faith both the subjective and objective elements of qualified immunity were satisfied. If the court meant the latter, which seems unlikely, the statement is not valid in light of the fact that there is no longer a requirement of subjective good faith for qualified immunity. Otherwise, an employee who violated clearly established case law would still be eligible for a statutory good faith defense merely by ignorance of the relevant case law (which is the cornerstone of the post-*Harlow* qualified immunity defense). Thus, even before the Supreme Court did away with the subjective aspect of qualified immunity, it does not seem that the statement in *Kilgore* was

---

92. Wolfel v. Sanborn, 666 F.2d 1005, 1007 (6th Cir. 1981), *vacated by* 458 U.S. 1102 (1982) (citing Patzig v. O'Neil, 577 F.2d 841, 848 (3d Cir. 1978)).

93. 18 U.S.C. §§ 2510-2522 (2000).

94. *Id.* § 2520.

95. Kilgore v. Mitchell, 623 F.2d 631, 633 (9th Cir. 1980) (citing Jacobsen v. Rose, 592 F.2d 515, 523 (9th Cir. 1978), *cert. denied*, 442 U.S. 930 (1979)).

96. *Id.* at 633-34 (citing Zweibon v. Mitchell, 516 F.2d 594, 671 (D.C. Cir. 1975), *cert. denied*, 425 U.S. 944 (1976)) (citing Wright v. Florida, 495 F.2d 1086, 1090 (5th Cir. 1974)).

97. *See* notes 31-35 and accompanying text.

meant to suggest that statutory good faith satisfied both aspects of qualified immunity analysis, but rather that a finding of statutory good faith satisfied the good faith requirement of the qualified immunity analysis.

If this hypothesis, that a finding of statutory good faith was sufficient for a finding of good faith under the pre-*Harlow* qualified immunity standard, is accepted, then the statement from *Kilgore* means that statutory good faith and the good faith question in qualified immunity analysis are roughly the same thing. Therefore, pre-*Harlow* good faith cases may provide insight to what a court would view as statutory good faith after the amendments to 19 U.S.C. § 482. However, such cases will not provide insight into what conduct is currently protected by qualified immunity. With this hypothesis in mind, it is beneficial to consult pre-*Harlow* decisions where qualified immunity was an issue to determine what constituted good faith.

The most useful cases for this analysis are those that roughly follow the theories that individuals assert against Customs employees. As stated, suits for violations of Fourth Amendment rights are brought against Customs employees under *Bivens*.[98] Suits against state employees for violation of federal civil rights are brought pursuant to 42 U.S.C. § 1983 ("§ 1983"). For purposes of qualified immunity analysis, suits brought under *Bivens* and § 1983 are treated similarly.[99] In fact, the rule seems to be that the defenses available to law enforcement officers under either theory are the same.[100] The notion that the defense of good faith applies to both *Bivens* and § 1983 cases is summed up in the following quote from *Brubaker*:

> The test, thus, under § 1983 is not whether the arrest was constitutional or unconstitutional or whether it was made with or without probable cause, but whether the officer believed in good faith that the arrest was made with probable cause and whether that belief was reasonable. It is now clear that an identical standard is to be applied in civil rights claims against federal officials based on the Fourth Amendment.[101]

Now that it is clearly established that the good faith standard applied to § 1983 cases is also applicable to *Bivens* cases, the question becomes what is good faith in either context. It is also beyond dispute that whatever the level of good faith that is required, the belief that the conduct was reasonable is essential to

---

98. Bivens v. Six Unknown Named Agents of the Fed. Bureau of Narcotics, 403 U.S. 388 (1999).

99. "[F]or the most part, courts have applied [§] 1983 law to Bivens [sic] cases." Rodriguez v. Ritchey, 539 F.2d 394, 399 (5th Cir. 1976) (citing Brubaker v. King, 505 F.2d 534 (7th Cir. 1974)).

100.

> We need not decide whether, under the facts of this case, it was appropriate to proceed against the federal defendants on the § 1983 theory since we are convinced that the standard for what constitutes a defense for a law enforcement officer is identical under § 1983 and the Fourth Amendment.

Brubaker, 505 F.2d at 536.

101. *Id.* at 536-37.

avoiding liability.[102] Finally, the defense of qualified immunity must be pled and proven by the defendant.[103] It is reasonable to assume that the same is true for a purely good faith defense. Some courts have shirked the question of what defines good faith. For instance, in *Procunier* the Supreme Court did not define what level of conduct would evidence a lack of good faith, but instead simply stated that the conduct in that case did not rise to the requisite level of malice, hinting that negligence in and of itself was not enough to deny the defense.[104] However, this reasoning is useful because it indicates that some sort of intentional injury is required.

Under this definition, good faith appears to merely be the absence of bad faith, with negligence not negating the protections offered by the defense. In fact, in *Shelton*, the court noted that there was "no evidence that [the Customs agents] acted other than in good faith."[105] Other courts have honed in on this reasoning and identified the difficult fact issues to which it gives rise. In *Putman* the court noted that reasonable minds might differ as to whether the conduct of the officer involved demonstrated a reasonable use of force to prevent a prisoner from escaping, or an unlawful and malicious intent to cause injury.[106] This argument again identifies the fact that the question of good faith is often not amenable to a summary judgment motion.

While good faith is often a fact based inquiry, some plaintiffs have argued that good faith is not available where the mistake was a mistake of law rather than a mistake of fact. Under this logic, an officer who made an arrest because of a misunderstanding of the law would not be entitled to immunity for his actions, and if the facts of the case were not in dispute, then the plaintiff may actually be able to prevail in early stages of litigation.

This argument was made in *Benson*, a case in which Customs agents executed a warrant based on probable cause that the plaintiffs were smuggling South African gold pieces (Krugerrands) without declaring them.[107] As it turned

---

102. Wood v. Strickland, 430 U.S. 308 (1974).

> The official himself must be acting sincerely and with a belief that he is doing right, but an act violating a student's constitutional rights can be no more justified by ignorance or disregard of settled, indisputable law on the part of one entrusted with supervision of students' daily lives than by the presence of actual malice.

Id. at 321.

103. Landrum v. Moats, 576 F.2d 1320, 1329 (8th Cir. 1978).

104. The Court stated that liability was authorized

> where the official has acted with "malicious intention" to deprive the plaintiff of a constitutional right or to cause him "other injury." This part of the rule speaks of "intentional injury," contemplating that the actor intends the consequences of his conduct . . . To the extent that a malicious intent to harm is a ground for denying immunity, that consideration is clearly not implicated by the negligence claim now before us.

Procunier v. Navarette, 434 U.S. 555, 566 (1978).

105. Shelton v. U.S. Customs Serv., 565 F.2d 1140, 1142 (9th Cir. 1977).

106. Putman v. Gerloff, 639 F.2d 415, 422 (8th Cir. 1981).

107. Benson v. Hightower, 633 F.2d 869 (9th Cir. 1980).

out, the Krugerrands were not subject to a requirement of declaration to Customs because they were currency. Due to this mistake the plaintiffs brought an action for violation of their Fourth Amendment rights.[108] The court did not dispute that the agents were entitled to qualified immunity, but the plaintiff argued that qualified immunity was not available where the agent's mistake was a mistake of law (the mistake being that the Krugerrands were declarable items).[109] The plaintiffs based their argument on the Restatement (Second) of Torts, which states in relevant part "no protection is given to a peace officer, who however reasonably, acts under a mistake of law other than a mistake as to the validity of a statute or ordinance."[110] Luckily for the Customs agents, the court was not persuaded by the logic of the Restatement. While the court did not hold that qualified immunity is always available when there is a mistake of law, it did hold that on the particular facts, including the fact that the arrest had followed after a successful search with a warrant which indicated that there had been a showing of probable cause by a magistrate, the justifications for applying qualified immunity were present.[111] This case again illustrates how strongly the courts have opposed attempts to chip away at protections offered to those in the public service. The court recognized that where the mistake was one of law, the plaintiff might have a valid argument that qualified immunity should not apply, but that argument was dependent on the facts of the case.[112]

While *Harlow* did away with the good faith aspect of qualified immunity because it dealt with objective factors, it did not hold that the defense was purely objective.[113] In fact, courts have held that the good faith defense can apply to reliance on regulations,[114] and even to operating procedures of local police departments.[115] Some courts have held, under certain statutory schemes, that subjective good faith alone does not suffice for a showing of good faith.[116] For instance, to show good faith under the Fair Labor Standards Act the official must

---

108. *Id.* at 870.

109. *Id.*

110. *Id.* (quoting RESTATEMENT (SECOND) OF TORTS § 121 cmt. i (1965)).

111. *Id.* at 871.

112. *Id.*

113. "Although the *Harlow* Court indicated that the good-faith defense turns *primarily* on objective factors, it did not hold that an *exclusively* objective standard was to be applied to claims that proceeded to trial. Thus, the County's argument that the standard for good-faith immunity must be purely objective, is untenable. . . ." Vizbaras v. Prieber, 761 F.2d 1013, 1016 (4th Cir. 1985) (quoting McElveen v. County of Prince William, 725 F.2d 954, 957-58 (4th Cir. 1984)).

114. Reliance on a policy that is later invalidated by a court is still grounds for good faith as long as the reliance occurred before the policy was invalidated. Kilgore v. Mitchell, 623 F.2d 631, 635 (9th Cir. 1980).

115. "Police officers also have a good faith reliance on standard operating procedures, if at the time of the incident they relied on the standard operating procedures of their institution when responding to an incident and when such reliance is honest and [sic] in its intention and reasonable." *Vizbaras*, 761 F.2d at 1015 (quoting jury instructions).

116. Int'l Ass'n of Firefighters, Local 349 v. City of Rome, Ga., 682 F. Supp. 522, 532 (N.D. Ga. 1988).

show that he acted in reliance on a written agency interpretation.[117] This requirement is not a creature of the courts, but is actually a requirement in the statute.[118]

Sometimes officials find that offering proof that they relied on agency procedures is the easiest way to show that they acted in good faith. Courts have proven receptive to such arguments, and it seems to be one of the most objective arguments available for proving good faith. In one case, officers were able to avoid liability for shooting an unarmed man because they relied on department procedure.[119] Under the revisions of 19 U.S.C. § 482 Customs inspectors will likewise be able to argue that they relied on their department procedures to avoid liability.

A related argument that defendants have made in good faith cases is that the defendant relied on orders from a superior. In a case in which a sheriff's deputy made a similar argument, the court stated that "[T]he fact that the actions are taken pursuant to orders and instructions is not a defense in and of itself, although it may be relevant to a claim of good faith and the defense of qualified immunity."[120]

In addition to looking at pre-*Harlow* qualified immunity cases, it is also beneficial to look at other statutory schemes that provide a good faith defense. While an in-depth analysis of these other schemes is not necessary, it is useful to consider whether the courts utilize a different analysis than under a qualified immunity standard. Title III of the Omnibus Crime Control and Safe Streets Act provides a good faith defense when a law enforcement agent acts in good faith and reliance on a court order or legislative authorization.[121] Cases, such as *Burkhart v. Saxbe*, that arise under this act often involve placement of wiretaps since that is one avenue that the act allows for in protecting the public.[122] This case was brought after the plaintiff's conversations were overheard during warrantless surveillance. Attorney General John Mitchell authorized the wiretaps without judicial authorization.[123] The plaintiffs contended that the wiretaps (which were not on their phones but on the phones to which the plaintiffs made calls) violated their Fourth Amendment rights and Title III of the

---

117. *Id.*

118. 29 U.S.C. § 259 (2000).

119. Landrum v. Moats, 576 F.2d 1320 (8th Cir. 1978). In this case the officers justified their use of deadly force on a directive in their department giving them the authority to use a firearm to aid the arrest or capture, and to prevent the escape, of a person who the officer has reasonable grounds to believe has committed a felony. *Id.* at 1323.

120. Putman v. Gerloff, 639 F.2d 415, 422-23 (8th Cir. 1981). The court cited *Forsyth* for the proposition that agents acting in good faith while following instructions would not be liable, but that agents that knew or should have known that they were violating the rights of a plaintiff could not "hide behind the cloak of institutional loyalty." *Id.* at 423 (quoting Forsyth v. Kleindienst, 599 F.2d 1203, 1217 (3d Cir. 1979)).

121. 18 U.S.C. § 2520 (2000).

122. 448 F. Supp. 588 (E.D. Pa. 1978).

123. *Id.* at 591.

Act.[124] The defendants responded that the wiretaps were authorized by the Attorney General and were exempt from Title III and from a requirement of prior judicial approval.[125]

The standard for good faith relied on by the court was that the defendants reasonably did not know that conducting the surveillance without a warrant violated the Fourth Amendment and that the agents acted without malicious intention to deprive the plaintiffs of constitutional rights, or to cause them to suffer some other injury.[126] The court went on to note that negligent interference with a plaintiff's rights was not a basis for denying the defense because there was no "intentional injury" to the plaintiffs.[127] Given these fact intensive standards, it is not surprising that the motions for summary judgment were denied so factual issues could be resolved.[128]

V. GOOD FAITH UNDER 19 U.S.C. § 482(b)

The task now arises of putting these conflicting standards into some order to roughly determine what the good faith defense provided by 19 U.S.C. § 482(b) really means. A few seemingly universal guidelines arise out of the preceding sections. Procedurally, the defendant has the burden of proving all elements of the defense to the satisfaction of the jury.[129] There must be some showing of an intentional injury. Mere negligence does not suffice to overcome the defense of good faith.[130] Also, the defendant's belief that his conduct was not unconstitutional must be reasonably held.[131] Finally, the defendant will likely be allowed to plead reliance on agency regulations as a defense to liability.[132]

As to the requirement that the belief be reasonably held, it has already been illustrated that both an objective and subjective showing of reasonableness is required, and that ignorance of the applicable constitutional standards is no defense to liability.[133] Therefore, the courts seem to be saying that Customs officers will be charged with knowledge of existing established constitutional principles when making determinations as to the reasonableness of inspectors' actions. If those actions violate such well established law, it appears that the inspectors will not be entitled to protection offered by the defense even if they were not aware of the law. On this level, the defense of good faith mirrors the post-*Harlow* analysis of qualified immunity—namely that the defendant will be held liable if his conduct violates established constitutional principles of which

---

124. *Id.*
125. *Id.* at 592.
126. *Id.* at 608-09 (citing Skehan v. Bd. of Trs., 538 F.2d 53, 62 (3d Cir. 1976) (en banc)).
127. *Id.* at 609-10 (citing Procunier v. Navarette, 434 U.S. 555, 556 (1978)).
128. *Id.* at 610.
129. Landrum v. Moats, 576 F.2d 1320, 1329 (8th Cir. 1978).
130. *See supra* text accompanying note 104.
131. Brubaker v. King, 505 F.2d 534, 536-37 (7th Cir. 1974).
132. *See supra* note 119 and accompanying text
133. Wood v. Strickland, 420 U.S. 308, 321 (1975).

he knew or reasonably should have known.[134] Good faith immunity will allow other means for Customs employees to avoid liability for their actions, but similar to qualified immunity, it will not allow them to avoid liability for actions that they knew or should have known were illegal.

The fact that good faith and qualified immunity defenses both attribute knowledge of well established constitutional principles to the defendant will largely negate the aspect of the good faith defense that would allow a defendant to avoid liability for actions that he or she truly believed, although erroneously so, to be valid. However, there are other benefits to the good faith defense that will still allow employees to avoid liability even when qualified immunity may not. Perhaps the most important of these means is that Customs employees will be able to prove good faith by reliance on agency procedures. In fact, because the good faith defense will likely have an objective component, reliance on agency procedures will likely be the most common means that Customs employees attempt to avoid liability. In the case of Customs, these procedures will likely come out of the *Personal Search Handbook* of the Customs Service.[135] In fact, throughout many of the existing cases against Customs employees, the employees cite their reasoning for conducting inspections to aid in their defense.[136] While it is unclear which of these reasons are specifically listed in the *Personal Search Handbook*, it is clear that courts do consider them as justifications for the decision to conduct a search. The result of this is that Customs inspectors will likely be able to plead and prove good faith by reliance on custom as well as written agency procedure. Of course there is no requirement that courts accept all proffered justifications for a search as establishing good faith when those justifications are not in writing,[137] but when the procedures are in writing, it seems likely that courts will not hold inspectors liable for conduct that conforms to that policy.

---

134. *See supra* note 32.

135. *See* Bradley v. United States, 299 F.3d 197, 204 (3d Cir. 1985) (discussing probable cause requirements contained in the *Personal Search Handbook*).

136. *See* Montoya de Hernandez v. Hernandez, 473 U.S. 531, 533-34 (1985) (citing such facts that plaintiff had arrived from a "source city" for narcotics, plaintiff spoke no English, and plaintiff carried large quantity of cash, plaintiff had only one pair of shoes); *see also* Brent v. Ashley, 247 F.3d 1294, 1297 n.1 (11th Cir. 2001) (justifying a strip search and x-ray that plaintiff fit the profile of African-American women on the same flight as Nigerian men, that she had arrived from a "source country," that she disapproved of the treatment of the only other black passenger on the flight, that the ticket had been purchased by a friend with a credit card at the same travel agency where the only other black passenger had purchased his ticket, she wore inexpensive clothes, and she became nervous and agitated when confronted).

137. *See Brent*, 247 F.3d at 1299 (stating that the fact that the inspectors saw the plaintiff shake her head in disapproval of the treatment of another black passenger was not sufficient to find reasonable suspicion to conduct a search).

VI. THE DEBATE

More can be said about what will constitute good faith in Customs inspection cases. Hypotheticals could be offered to illustrate when the good faith defense may protect inspectors in situations where qualified immunity fails to, or vise versa. However, these are issues that will be addressed by courts over time, using the framework and reasoning already discussed in this note. What remains then is not to better understand how the good faith defense will work, but why it is necessary at all. Obviously, there must have been some concern when the Trade Act was passed that led Congress to take such drastic action. In the remainder of this Note I will discuss both positions for and against the addition of this layer of protections for Customs employees, and attempt to balance those interests to see if the defense is a legitimate expansion of a necessary protection or overreaching legislation capitalizing on the tragedy of September 11, 2001.

To begin, it is useful to understand the context that suits against Customs officers will be brought. Customs inspectors search incoming shipments of merchandise at sea ports. They also search people and merchandise at inland ports of entry on the Canadian and Mexican borders as well as international airports. Because inspections at ports of entry include people and not just merchandise, and because there is no requirement of probable cause for normal searches conducted at the borders,[138] the issue of racial profiling will often come up in the context of Customs searches.

The possibility that racial profiling may occur, coupled with the possibility that an inspector guilty of profiling may be able to avoid liability by claiming he or she acted in good faith, is a strong argument against the addition of the defense. To make a claim for violation of equal protection rights in the racial profiling context, the claimant will have to prove that the actions taken by government officials had a discriminatory effect, and were motivated by a discriminatory purpose.[139] In order to prove discriminatory effect the claimant has the burden of showing that he or she is a member of a protected class and that he or she was treated differently from similarly situated individuals in an unprotected class.[140]

One can see the difficulties facing a claimant attempting to make a charge of racial profiling against Customs in the context of a search, in order to prove that he or she was treated differently from other members of a protected class. The plaintiff will not have access to this information short of extensive discovery, which is very costly to perform. Many individuals who feel their rights have been violated may not have the means to hire a lawyer and look into allegations of wrongdoing, much less the means to pay for extensive discovery. Further complicating the task of a claimant attempting to protect their civil rights is the fact that qualified immunity often arises in the early stages of litigation.

---

138. *See supra* note 52.

139. *Bradley*, 299 F.3d at 205 (citing Arlington Heights v. Metro. Hous. Dev. Corp., 429 U.S. 252, 264-66 (1977)) (citing Washington v. Davis, 426 U.S. 229, 239-42 (1976)) (citing Chavez v. Illinois State Police, 251 F.3d 612, 635-36 (7th Cir. 2001)).

140. *Id.* at 206 (citing *Chavez*, 251 F.3d 612 at 636).

Therefore, Customs may be able to win on a motion for summary judgment before the claimant even knows what discovery he or she must conduct. This is exactly what happened in *Bradley*. There, the plaintiff alleged that the district court restricted discovery and then granted summary judgment even though further discovery was necessary to her case. The circuit court upheld the ruling of the lower court, noting that Bradley had the responsibility under the Federal Rules of Civil Procedure, Rule 56 (f), to file an affidavit that specifically identified the information sought, how that information would preclude summary judgment, and why it had not been obtained earlier.[141] This argument is not presented to suggest that the plaintiff's failure to follow procedural rules should be excused. Instead, it is offered as one more illustration of how high the burden is on a plaintiff complaining against the Customs Service. Not only is it difficult to obtain a favorable ruling against the government, but also it is difficult to be able to make a claim against the government given the procedural safeguards in place. Of course, with the passage of the good faith defense, even if a plaintiff is able to successfully survive the discovery and summary judgment stages of litigation, claims for racial profiling will still be difficult to win given the fact that inspectors will be able to plead reliance on agency procedures as a defense.

In support of the defense, there is no question that some protections must be afforded to the government in order to avoid unchecked litigation and frivolous lawsuits. The government must be able to function and perform its duties to the public in general without fear of liability from lawsuits filed by overly litigious citizens, and without fear of defending frivolous lawsuits beyond the earliest stages of litigation. In fact, the societal costs of extended litigation to the public, and the resulting loss to society of productive use of governmental assets, are among the very justifications put forth for the qualified immunity defense in *Harlow*.[142] However, the addition of a good faith defense to Customs' arsenal of defenses to liability flies directly in the face of concerns regarding the amount of litigation the government must contend with, and is not justified by any failure of the existing defenses to adequately protect Customs employees from liability.

It has been illustrated that the question of good faith will require determinations of fact in most cases where qualified immunity fails to protect the inspector, especially those where the inspector is not found to act in good faith on reliance of agency procedures. It is clear that the question of good faith will not come into play until the motion for summary judgment has been denied. Once an inspector's claim of qualified immunity is denied on summary judgment, the inspector still has the opportunity to defend the lawsuit on the merits. Only at this stage of litigation, when the inspector is forced to defend the lawsuit on the merits, will the issue of good faith begin to be litigated. Therefore, arguments that society's resources should not be wasted in litigating frivolous lawsuits do not support passage of the good faith defense because lawsuits where good faith is at issue are ultimately going to be litigated. In effect, the good faith defense simply provides another method of avoiding liability during the trial without serving the goals of qualified immunity, which include conservation of

---

141. *Id.* at 206-07.

142. *See supra text* accompanying note 31.

government resources and allowing the government to function efficiently.

In three recent cases involving claims against Customs inspectors, qualified immunity failed to protect the inspector in only one instance. In *Saffell v. Crews*, inspectors were not held liable for a partial strip search because a bulge in the claimant's pelvic region was grounds for reasonable suspicion of contraband.[143] No drugs were found on the plaintiff, but the court still found that the inspectors were entitled to qualified immunity for their actions. As discussed in *Bradley*, the inspectors were not held liable because their conduct was held not to be the subject of a requirement of reasonable suspicion.[144] Only in *Brent* were the inspectors denied the protection of qualified immunity.[145]

It is therefore clear that qualified immunity is not failing to meet its goals of both protecting the civil rights of citizens and protecting the government (specifically Customs) from frivolous lawsuits. What then is a reasonable explanation of why more protections were required for the inspectors in these cases? One possible explanation is that in a knee-jerk reaction to September 11, 2001, public concern over the protection of the nation's borders was high enough that this provision passed Congress without much concern over the possibility that it would be abused. Another possibility is that the amendment passed simply because it was buried in a massive Trade Act. Indeed, because of the sheer size of the Trade Act of 2002, debate over the specific amendment to 19 U.S.C. § 482 is scarce, but what is there is very revealing of the concerns attendant with the legislation. Numerous civil rights organizations, including the NAACP, Leadership Conference on Civil Rights, ACLU, and the Council on American Islamic Relations all expressed strong opposition to passage of the amendment to § 482.[146] Beyond special interest groups, a number of members of Congress expressed concern over the provisions of the Trade Act dealing with Customs immunity. Perhaps the most compelling statements against the passage of the amendment dealt with a report by the General Accounting Office ("GAO"). This report, which focused on racial profiling, found that African-American women were nine times more likely than Caucasian women to be the subjects of intrusive searches while they were only half as likely as Caucasian women to be found carrying contraband.[147]

Some representatives focused on the obvious need to protect national security, but emphasized that this need should not come at the expense of civil rights.[148] One congressman specifically emphasized that the amendment was a capitalization on the tragedy of September 11, 2001, which was being used to justify offenses to the Constitution.[149] The same congressman went on to note that Customs had not provided any justification why qualified immunity was not

---

143. *See supra* note 3 and accompanying text.

144. *See supra* note 70 and accompanying text.

145. *See supra* notes 74-85 and accompanying text.

146. 148 CONG. REC. H5980-H5982 (daily ed. July 26, 2002).

147. *Id.* at H5978-H5979 (statement of Rep. Jackson-Lee).

148. *Id.* at H5979.

149. *Id.* at H5979 (statement of Rep. Conyers).

protection enough, especially in light of the GAO's study.[150]

Aside from concerns about the amendment, the statements of the representatives during the debates over the Act provide some insight into how the members of Congress understood the good faith defense would operate. There was concern that an officer "could engage in blatantly discriminatory conduct, but if he in 'good faith' believed that he was justified in doing so, he could not be held liable," and that a claimant would be entitled to no relief unless the inspector acted in "bad faith."[151] There was also concern about how the proposal would affect the population.[152] Unfortunately, based on the analysis of what constitutes good faith in the law enforcement context, these statements are entirely accurate as to what the effect of this legislation will likely be.

The arguments made in Congress sufficiently state the issue. No clear reason exists to justify giving Customs officials more immunity than any other law enforcement agency. Good faith was originally a factor in the qualified immunity analysis that the Supreme Court later did away with in *Harlow* because it would waste too many government resources in litigation. Now Customs employees have the benefits of good faith immunity, possibly even if they reasonably should have known that their actions violated clearly established law, without the benefits to society that stem from qualified immunity. Given the past holdings in cases dealing with challenges to searches conducted under the authority of 19 U.S.C. § 482, there is no showing of necessity for this extra layer of protection. Obviously, the events of September 11, 2001, and the resulting war on terror have increased the public's concern about who and what are entering the country, but should rights that have been fundamental since the adoption of the Constitution be sacrificed in a knee-jerk reaction to those events? Given the delicate balance that courts have sought to achieve since the adoption of qualified immunity, and the lack of sound policy supporting the additional protections afforded to Customs in this instance, it is difficult to see this legislation as anything but a reflex to September 11, 2001, that may lead to unintended consequences.

---

150. *Id.*

151. 148 CONG. REC. H5979 (daily ed. July 26, 2002).

152. "This proposal would hurt real people. It would increase the likelihood of meritorious claims being thrown out. Parties would end up fighting at length over whether an official did or did not subjectively believe his conduct to be lawful—even if existing law clearly established that it wasn't." *Id.* at H5980.

DEPECAGE: EMBRACING COMPLEXITY TO SOLVE CHOICE-OF-LAW ISSUES

CHRISTOPHER G. STEVENSON*

INTRODUCTION

An airplane crashes in the Arizona desert with no survivors. There are numerous causal links to the crash, such as: (1) a defective aircraft engine manufactured in Missouri by a company incorporated in Delaware; (2) a defectively designed airframe made in France by a company that has its U.S. headquarters in Virginia; (3) negligent actions of the pilot who was employed by a Texas company, but was trained and certified by a Florida based company through a contract with the pilot's employer; (4) negligent actions of the Federal Aviation Administration (FAA) air traffic controller in Nevada; (5) faulty information on the pilot's aeronautical maps, which were printed by a governmental agency in Washington D.C.; and (6) poor weather conditions which were not reported by an FAA weather observer in New Mexico. Additionally, there were passengers on board from twenty different states. The question is: Which state's law applies?

A difficult hypothetical question to say the least, but not far from the truth. Courts have been faced with factual scenarios similar to this hypothetical not only in aircraft disasters, but also in complex mass tort litigation where tortfeasors and victims are dispersed across the country.[1]

Then, what does a court do when it faces a situation similar to the one posed under this hypothetical scenario? Does a court pick one state's law and apply it to all parties on all issues? By using this approach a court could, devoid of any legal analysis, find that because the plane crashed in Arizona, Arizona law should apply across the board. On the other hand, a court could choose to apply different states' laws depending on a state's relationship to individual defects, negligent acts, passenger's domicile or residency in determining liability and damages. Under the hypothetical, the court might apply Missouri law to the engine manufacturer who produced the defect in the doomed aircraft's engine, but apply Virginia law as to any defects in the airframe. Furthermore, the court could calculate compensatory damages for passenger Smith, a Californian, according to California law, and for passenger Jones, from Georgia, according to Georgia law.

It is this approach of applying various states' laws to separate issues which has been labeled as "depecage." Specifically, "under the doctrine of depecage,

---

* J.D. Candidate, 2004, Indiana University School of Law—Indianapolis; B.S., 1996, Purdue University.

1. Use of depecage in mass tort litigation is frequently found in large class action suits. A class action suit filed against Philip Morris, Inc. provided fertile ground for the court's application of depecage. *See* Simon v. Philip Morris, Inc., 124 F. Supp. 2d 46, 75-78 (E.D.N.Y. 2000); *In re* Simon II Litigation, 211 F.R.D. 86, at 111 (E.D.N.Y 2002) (applying principles of depecage to multi-state tobacco litigation). *But see In re* Bridgestone/Firestone, Inc., 288 F.3d 1012 (7th Cir. 2002) (failing to mention depecage as a possible solution).

different substantive issues in a tort case may be resolved under the laws of different states where the choices influencing decisions differ."[2] This Note attempts to explore depecage and offer it as a needed tool for efficient and effective choice-of-law analysis. Part I provides examination of the origins of depecage from modern day choice-of-law principals. Part II discusses the mechanics of applying depecage to complex litigation, and demonstrate how various courts have used depecage as a successful judicial instrument.

Part III explores the reasons why some courts have not embraced depecage. There are some theoretical underpinnings behind this counter-rationale, and a tremendous push by some legal scholars to disregard the doctrine completely. Assisting in this counter-movement against depecage are those cases where either the court used depecage without any formal recognition of what it was doing, or failed to apply depecage to a situation which desperately called for its use.[3] Part IV discusses the lack of direction that has left courts in confusion with little knowledge and guidance from higher authority on the application of depecage to choice-of-law problems. While most courts have not expressly rejected depecage, their use of the doctrine has hardly been a model of clarity.

Despite the opposition to the doctrine expressed by courts and the indifference exhibited by others, Part V provides the rationale for applying depecage and the necessity of its application to an otherwise irreconcilable conflict.[4] Part V contends that, when dealing with complicated choice-of-law issues, courts should actively embrace complexity by applying the doctrine of depecage. Depecage allows courts to isolate and limit true conflicts between differing bodies of law, which facilitates more adequate analysis of underlying interests and policies.[5] By promoting diversity among state law, depecage is also consistent with the federalist view of state sovereignty.[6] Thus, through its use, courts will provide better direction for the doctrine, protect state and individual interests, and eventually broaden its use.[7]

I. THE ORIGINS OF DEPECAGE

A French word, dépeçage (DE-PA-SAJ) is defined as a "cutting up, dismembering, carving up."[8] In a legal setting "depecage" is "[a] court's application of different state laws to different issues in a legal dispute; choice of

---

2. LaPlante v. Am. Honda Motor Co. Inc., 27 F.3d 731, 741 (1st Cir. 1994).

3. *See* Allen v. Great Am. Reserve Ins. Co., 766 N.E.2d 1157 (Ind. 2002) (applying Indiana law to the construction of a contract and South Carolina law to claim based on multistate agency principle); *see* discussion *infra* Part III.

4. *See* Schalliol v. Fare, 206 F. Supp. 2d 689, 700 n.30 (E.D. Pa. 2002).

5. *See infra* Part V.A.

6. *See infra* Part V.B.

7. Courtland H. Peterson, *Private International Law at the End of the Twentieth Century: Progress or Regress?*, 46 AM. J. COMP. L. 197, 224 (1998).

8. COLLINS ROBERT FRENCH ENGLISH DICTIONARY 233 (4th ed. 1995).

law on an issue-by-issue basis."[9] In other words, depecage is "the process of cutting up a case into individual issues, each subject to a separate choice-of-law analysis."[10]

A. Vested Rights and Lex Loci Delicti

Before the modern choice-of-law approach used by courts which entertains concepts of a party's relationship or contact with other states, courts strictly enforced the doctrine of *lex loci delicti.* Under this "vested rights" doctrine, in determining which state's law to apply, a court looks solely at the place where the tort was committed.[11] This doctrine was in full force at the turn of the century when choice-of-law questions were often raised by railroad accidents.

In *Northern Pacific Railroad Co. v. Babcock*, a locomotive engineer was killed when the train derailed in Montana.[12] His estate brought suit in Minnesota citing Montana law; however, the defendant wanted Minnesota law to apply.[13] The defendant admitted the general rule of *lex loci delicti*, but contended that, because plaintiff's action was founded in state statutory language, the law of the forum and the law of the place where the right of action accrued must concur in order for the *lex loci* to apply.[14] Quoting from the Supreme Court of Minnesota, the Court reasoned that:

> [I]t by no means follows that, because the statute of one state differs from the law of another state, therefore it would be held contrary to the policy of the laws of the latter state. Every day our courts are enforcing rights under foreign contracts where the *lex loci contractus* and the *lex fori* are altogether different, and yet we construe these contracts and enforce rights under them according to their force and effect under the laws of the state where made.[15]

Under *lex loci delicti*, the Court held that because the contract of employment was made in Montana, and the accident occurred in that state, there was no error in holding that the right to recover was governed by Montana law.[16]

Justice Holmes strongly affirmed the doctrine of *lex loci delicti* in *Slater v. Mexican National Railroad Co.*[17] In *Slater*, a Texas resident, employed by a Colorado railroad company, was killed while coupling cars together in Nueva

---

9. BLACK'S LAW DICTIONARY 448 (7th ed. 1999).

10. *See* Ruiz v. Blentech Corp., 89 F.3d 320, 324 n.1 (7th Cir. 1996).

11. BLACK'S LAW DICTIONARY, *supra* note 9, at 923.

12. N. Pac. R.R. Co. v. Babcock, 154 U.S. 190, 196 (1894) (quoting Harrick v. Minneapolis & St. Louis Ry. Co., 16 N.W. 413, 414 (Minn. 1883)).

13. *Id.* at 197.

14. *Id.*

15. *Id.* at 198.

16. *Id.* at 199.

17. Slater v. Mexican Nat'l R.R. Co., 194 U.S. 120 (1904).

Laredo, Mexico. His widow and children brought suit in Texas.[18] Without any discussion of relationships, contacts, or state interests the Court stated that "[a]s the cause of action relied upon is one which is supposed to have arisen in Mexico, under Mexican laws, the place of the death and the domicile of the parties have no bearing upon the case."[19] Thus, the *Slater* Court affirmed the circuit court's ruling dismissing plaintiffs' claim brought in Texas.[20] This "vested rights" approach was enshrined in the 1934 Restatement (First) of Conflict of Laws, which advocated hard and fast choice-of-law rules "premised on the principle that the last event necessary to create or change a legal relationship determines where a right vests."[21]

However, even in 1894 under *National Pacific Railroad Co. v. Babcock*, the Court recognized that there may be some situations where the use of *lex loci delicti* should not be used. In order to justify the disregard of *lex loci delicti*, "it must appear that it is against good morals or natural justice, or that, for some other such reason, the enforcement of it [*lex loci delicti*] would be prejudicial to the general interests of our own citizens."[22] Such policy considerations had been argued by legal scholars for many years prior to their adoption by American courts. Scholars discredited *lex loci delicti* because it failed to take into account underlying policy considerations in evaluating the significance to be ascribed to the circumstance that an act had concerning the rights and liabilities which arise out of that act.[23]

B. The Move Toward Interest Analysis

As our nation began to industrialize, the mass production and marketing of goods such as telephones, radios, wire services, and automobiles, on a national scale, made the nation seem smaller and more closely knit.[24] This national industrialization corresponded with a diminished sense of the importance of state lines that served to divide people. In the mobile Twentieth Century, courts began to ask themselves whether the results dictated by the inflexible rules of *lex loci delicti* and vested rights retained their past attraction.[25] Courts began to decide cases in ways that made "good socio-economic sense."[26] However, the rule of

---

18. *Id.* at 124.

19. *Id.* at 127.

20. *Id.* at 126, 131.

21. Simon v. Philip Morris, Inc., 124 F. Supp. 2d 46, 65 (E.D.N.Y. 2000) (quoting J. BEALE, A TREATISE ON THE CONFLICT OF LAWS 1288 (1935)).

22. N. Pac. R.R. Co. v. Babcock, 154 U.S. 190, 198 (quoting Hervick v. Minneapolis & St. Louis Ry. Co., 16 N.W. 413 (Minn. 1883)).

23. Harold P. Southerland, *Sovereignty, Value Judgments, and Choice of Law*, 38 BRANDEIS L.J. 451, 478-79 (2000).

24. *Simon*, 124 F. Supp. 2d at 66; Southerland, *supra* note 23, at 471-72.

25. *Simon*, 124 F. Supp. 2d at 66; Babcock v. Jackson, 191 N.E.2d 279, 283 (N.Y. 1963); Southerland, *supra* note 23, at 471.

26. Southerland, *supra* note 23, at 473 (quoting Robert A. Lefler, *Conflict Laws: More on*

lex loci delicti still formally governed tort actions. To circumvent this rigid rule and give courts an excuse to look beyond the law of the state where the action accrued, parties began to invent escape devices.[27] One such escape device was the characterization of a tort claim as a contract claim, which raised the contention that the issue was more procedural than substantive.[28]

In the 1960s, courts began to change course and take into consideration varying factors separate from *lex loci delicti.* The watershed case of *Babcock v. Jackson* is a defining moment in choice-of-law analysis.[29] In *Babcock*, an automobile passenger from Rochester, New York, was seriously injured when the driver (also from Rochester) lost control of a car in Ontario, Canada.[30] The passenger, Ms. Babcock, filed suit in New York, where she had a cognizable claim. A Canadian statute absolved the driver of any liability for injuries sustained by a guest passenger in his or her vehicle.[31] The court resoundingly stated that the law of New York should apply regardless of where the accident occurred.[32] In so doing, the *Babcock* court thoroughly denounced the vested rights and *lex loci delicti* doctrine. The court stated that "the vice of the vested rights theory, it has been aptly stated, 'is that it affects to decide concrete cases upon generalities which do not state the practical considerations involved.'"[33] The court instead used a "center of gravity" or "grouping of contacts" to analyze the choice-of-law issue.[34] The court explained this new test as a way to achieve the best practical result "by giving controlling effect to the law of the jurisdiction which, because of its relationship or contact with the occurrence or the parties has the greatest concern with the specific issue raised in the litigation."[35] By drawing upon the ideas legal scholars had been advocating for years, the *Babcock* court laid the ground work for other courts to follow suit in denouncing *lex loci delicti.*[36]

However, *Babcock* left many questions regarding this new approach unanswered for other courts. The *Babcock* opinion furnished no guidance whatsoever for situations where all the contacts fail to converge in a single state. For example, what should the result be when a trip's starting point and destination are in different states, or codomiciliaries of one state rent a car

---

Choice-Influencing Considerations, 54 CAL. L. REV. 1584, 1587-88 (1966)).

27. *Id.*

28. *Id.*

29. *Babcock*, 191 N.E.2d at 279; Harold L. Korn, *The Choice-of-Law Revolution: A Critique*, 83 COLUM. L. REV. 772, 827 (1983).

30. *Babcock*, 191 N.E.2d at 280.

31. *Id.*

32. *Id.* at 285.

33. *Id.* at 281 (quoting Hassel E. Yatena, *The Hornbook Method and the Conflict of Laws*, 37 YALE L.J. 468, 482-83 (1928)).

34. *Id.* at 283.

35. *Id.*

36. Korn, *supra* note 29, at 831.

licensed, garaged and insured in another.[37] This same type of failure, where higher courts fail to give adequate guidance on choice-of-law analysis, continues to occur today.[38]

C. The Restatement (Second) of Conflict of Laws

After almost twenty years in the making, in 1971, the Restatement (Second) of Conflict of Laws ("Restatement") attempted to summarize the nature of choice of law in American courts under the new "significant relationship test." Section 6 of the Restatement introduced the various considerations of interest analysis as:

> (a) the needs of the interstate and international systems, (b) the relevant policies of the forum, (c) the relevant policies of other interested states and the relative interests of those states in the determination of the particular issue, (d) the protection of justified expectations, (e) the basic policies underlying the particular field of law, (f) certainty, predictability and uniformity of result, and (g) ease in the determination and application of the law to be applied.[39]

From the Restatement section 145 came the expressed language of the doctrine of depecage. Section 145 introduced this concept as:

> (1) The rights and liabilities of the parties with respect to an issue in tort are determined by the local law of the state which, **with respect to that issue**, has the most significant relationship to the occurrence and the parties under the principles stated in § 6.
> (2) Contacts to be taken into account in applying the principles of § 6 to determine the law applicable to an issue include:
> (a) the place where the injury occurred,
> (b) the place where the conduct causing the injury occurred,
> (c) the domicile, residence, nationality, place of incorporation and place of business of the parties, and
> (d) the place where the relationship, if any, between the parties is centered.
> These contacts are to be evaluated according to their relative importance **with respect to the particular issue**.[40]

Thus, by calling for choice-of-law analysis for each particular issue and by providing what contacts are to be considered, the Restatement lays a framework for the application of depecage.

II. MECHANICS OF USING DEPECAGE: WHICH COURTS HAVE USED THE DOCTRINE AND WHY

Choice-of-law analysis involving the use of depecage often involves complex

---

37. *Id.* at 835.
38. Allen v. Great Am. Reserve Ins. Co., 766 N.E.2d 1157 (Ind. 2002); *see infra* Part IV.
39. RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 6 (1971).
40. *Id.* § 145 (emphasis added).

legal and factual parameters. In order to better understand how these complex factors work together it is beneficial to carefully examine how courts, which have been particularly adept in applying depecage, have used the doctrine in deciding choice-of-law issues. However, before considering how courts have used depecage, it is important to understand how depecage works through the acknowledgment of its driving forces.

A. Driving Forces of Depecage

The legal concepts of federalism, advocacy, and the prevention of forum shopping create choice-of-law problems which may require the use of depecage. While these concepts do not promote a particular rationale for or against the usage of depecage, they do provide a honest look at why choice-of-law becomes an issue to be dealt with in the first place.

The constitutional concept of federalism and the individual sovereignty it conveys to each state cuts to the core of depecage. Federalism is defined as "[t]he doctrine holding that a federal court must refrain from hearing a constitutional challenge to state action if federal adjudication would be considered an improper intrusion into the state's right to enforce its own laws in its own courts."[41] Justice Marshall also described the concept of federalism in *M'Culloch v. State*. He stated "[i]n America, the powers of sovereignty are divided between the government of the Union, and those of the states. They are each sovereign, with respect to the objects committed to it, and neither sovereign, with respect to the objects committed to the other."[42]

Federal/state political structuring creates the phenomenon of differential state sovereignty.[43] When the federal government was created, certain powers previously subject only to state authority were vested in the federal government.[44] Because every state gave up the same powers to the federal government, states still possessed identical powers after the creation of the federal government.[45] While all the states possess the same authority, it was expected that the states will exercise these powers differently. "The states are allowed—indeed, expected—to disagree on substantive issues."[46] The most valuable aspects of federalism is what courts and commentators frequently have recognized as the fifty state laboratories, which provide for the development of new social, economic, and political ideas.[47] Additionally, federalism preserves the ability of citizens to learn the democratic processes through participation in local

---

41. BLACK'S LAW DICTIONARY, *supra* note 9, at 1128.

42. M'Culloch v. Sate, 17 U.S. 316, 410 (1819).

43. Lea Brilmayer & Ronald D. Lee, *State Sovereignty and the Two Faces of Federalism: A Comparative Study of Federal Jurisdiction and the Conflict of Laws*, 60 NOTRE DAME L. REV. 833, 852 (1985).

44. *Id.*

45. *Id.*

46. *Id.*

47. Fed. Energy Regulatory Comm'n, 456 U.S. 742, 788 (1982).

government and to govern their local problems.[48] Thus, federalism enhances states' ability to disagree with each other creating the atmosphere necessary for the use of depecage.

Federalism and its creation of state sovereignty contributes to choice-of-law problems because each state is free, and even encouraged, to develop its own sovereign body of law. However, courts pay a high price for federalism in the enormous challenge of respecting each state's sovereignty. Judicial resources, including time and effort, are spent attempting to respect each state's sovereignty in determining whose law will apply. Application of one federal body of law to all multi-state actions would conserve judicial resources.[49]

The second root of depecage and modern day choice of law is the concept of advocacy. "A lawyer shall act with reasonable diligence and promptness in representing a client."[50] Lawyers must diligently advocate for their clients. Carrying out this charge often calls for lawyers to contend that one state's law should apply over another. If there is a remote chance that a more liberal state's law will offer a victim more compensation, a plaintiff's lawyer will most often zealously pursue the application of that state's law. Similarly, a lawyer counseling the defendant will be equally adamant that a more conservative state's law should apply. Thus, when applicable, courts are faced with motions to apply whichever state's law best helps each party's position. If neither party cared which state's law should apply, then a court's job in choosing applicable law would be much easier.

The third driving force of depecage and modern day choice-of-law principles is the goal of preventing "forum shopping." The goal of *Erie Railroad Co. v. Tompkins*, in prescribing uniformity between federal and state courts, was to prevent forum shopping between those court systems.[51] While *Erie* deals with state and federal forums, the same danger of forum shopping occurs among states when a *lex fori* choice-of-law approach is used. If there were a uniform rule that the law of the forum governs a claim, the problem would arise that all injured victims would run to file their claim in a forum with the most favorable law for their situation. Thus, a flood of litigation would ensue in only certain forums. At the same time, refusing to apply the law of the forum to an entire cause of action in turn requires selecting which state's law will apply. Again, conflicts

---

48. *Id.* at 789-90.

49. The face of our political design may be changing with the enactment of the Multiparty, Multiforum Trial Jurisdiction Act, 28 U.S.C § 1369 (2003). The Act gives federal district courts original jurisdiction of any action involving "minimal diversity" between adverse parties from a single accident where at least seventy five persons have died. This Act directly attacks the *Erie* doctrine in that it will create federal common law, threatening federalism by arbitrarily preempting state law. Undoubtedly, the constitutional basis of this newly enacted law will be questioned, which may force the United State's Supreme Court to defend its era of federalism and finally speak on a conflicts of law issue.

50. MODEL RULES OF PROF'L CONDUCT R. 1.3 (1999).

51. Maryellen Corna, *Confusion and Dissension Surrounding the Venue Transfer Statutes*, 53 OHIO ST. L.J. 319, 328 (1992).

analysis would be simplified, if not eliminated, by the courts ability to apply the law of the forum to all issues.

After examining three driving forces—federalism, advocacy, and the prevention of forum shopping—one can quickly realize that these principles in American law are not likely to disappear any time in the near future. Therefore, some framework for choosing applicable law must be in place which makes depecage an essential part of this framework.

B. Air Crash Disaster Litigation: Fertile Ground for the Application of Depecage

Air crash disaster litigation has been the primary means by which the principles and application of depecage has grown. In *In re Disaster at Detroit Metropolitan Airport on August 16, 1987* ("*Detroit Metro*"), the court recognized the need for depecage in the choice-of-law problem presented.[52] The court found where there are legally significant facts which have occurred in more than one state, the court must identify those states that have sufficient contacts with the litigation. Once these states have been identified, the court must determine whether the various substantive laws at issue differ with regard to the particular issues in contest. In other words, is there a conflict? If so, this may result in the application of "the rules of different states to determine different issues in the same case."[53]

In *Detroit Metro*, the plaintiffs filed wrongful death claims against McDonnell Douglas Corporation and Northwest Airlines seeking compensatory and punitive damages.[54] In deciding to use depecage, the court used Michigan's choice-of-law rule for claims filed in Michigan.[55] The *Detroit Metro* court first split the plaintiffs' products liability claims against McDonnell Douglas from their punitive damages claim against Northwest Airlines.[56]

In determining which state's law to apply to the product liability claim against McDonnell Douglas, the court found three possible states: (1) Michigan—the place of the crash; (2) Missouri—McDonnell Douglas' principal place of business; and (3) California—the place of manufacture and design of the accident aircraft.[57] In *Detroit Metro*, there was a conflict between Michigan's

---

52. *In re* Disaster at Detroit Metro. Airport on Aug. 16, 1987, 750 F. Supp. 793, 796 (E.D. Mich. 1989) [hereinafter Detroit Metro].

53. *Id.* (quoting Willis L. McReese, *Depecage: A Common Phenomenon in Choice of Law*, 73 COLUM. L. REV. 58, 75 (1973)).

54. *Id.*

55. Although Michigan had abandoned the doctrine *lex loci delicti*, at that time, they had declined to adopt a particular choice-of-law rule. Without expressly adopting the Restatement (Second) of Conflict of Laws, *see supra* notes 39-40 and accompanying text, the court in *Detroit Metro* did hint that Michigan law may be displaced when there are particular interests that each state has in having its substantive law apply to the precise issue in question. *Id.* at 797.

56. *Id.* at 799, 804-05.

57. *Id.* at 801-02.

products liability law, which applied a negligence standard to design defect claims, and Missouri and California products law which used strict liability.[58] The court analyzed which state's rule of law had the greater interest in seeing that its law was applied. The court discounted Michigan's interest stating that "Michigan simply has no interest in applying its law to protect a foreign state producer that supplies products for a company doing business in that state."[59] In contrast, the court found that both Missouri and California had a strong interest in applying their products liability law. The court reasoned that Missouri's and California's strict liability laws "reflect a desire to (1) regulate culpable conduct occurring within its borders, (2) induce corporations to design safe products and deter future misconduct, and (3) impose the financial repercussions, which have been incurred by the user of a defective product, upon the producer."[60] Because Missouri's and California's product liability laws were essentially identical, the court could have used either state's law, but the court ultimately concluded that California law should be applied since the alleged wrongful conduct occurred within its borders.[61] Thus, after splitting out the issue of product liability, the *Detroit Metro* court found a "false conflict" existed between Michigan's interest in seeing its law applied and California's interest.

With regard to the punitive damage claim against Northwest Airlines, the court decided that either Michigan law (the place of injury) or Minnesota law (the place of alleged misconduct and principal place of business) would apply.[62] The court found a conflict between the law of Michigan and Minnesota. Minnesota law allowed for punitive damages, while Michigan law did not.[63] As far as state interests in seeing its law applied to this particular issue, Michigan had a strong interest in prohibiting the imposition of punitive damages on companies doing substantial business within its borders.[64] On the other hand, Minnesota, Northwest's principal place of business, had an interest in deterring and preventing any wrongful conduct by corporations which locate their corporate headquarters within its borders.[65] This presented a "true conflict" for the court because application of either Minnesota's law or Michigan's law would undoubtedly undermine the policy which is reflected in the law of the other state. The court is then back to square one, and must make an arbitrary decision not unlike a *lex loci delicti* decision. In so doing, the court held that, as the forum state, the law of Michigan (which would not allow punitive damages) should apply.[66]

The series of decisions written by Judge Ruben Castillo concerning choice

---

58. *Id.* at 799-800.
59. *Id.* at 801.
60. *Id.* at 802.
61. *Id.*
62. *Id.* at 804-06.
63. *Id.* at 805-06.
64. *Id.* at 807.
65. *Id.*
66. *Id.* at 808 (applying the law of the forum where a true conflict is presented).

of law in the air crash near Roselawn, Indiana utilized depecage as well. In *In re Air Crash Disaster Near Roselawn, Ind. on Oct. 31, 1994*, the court found importance in understanding that

> the search for the applicable law is not a general one, but rather it is one that takes proper notice of the fact that the significance of a state's relationship to a particular aviation disaster may vary as a function of the particular issue presented. Consequently, under the doctrine of depecage, it is not uncommon for courts to apply the substantive law of several different states in resolving air crash cases.[67]

Roselawn involved flight 4184, which crashed in Indiana en route from Indianapolis to Chicago. Flight 4184 was an Avions de Transport Regionale (ATR) aircraft built in France.[68] The day of the accident, 4184 had touched four different states and at the time of the accident was on its final leg of the day.[69] Plaintiffs claimed that both American Airlines and ATR were liable for the ice accumulation on the aircraft's wings which ultimately led to the plane's instability, loss of control, and subsequent destruction, killing all on-board.[70]

Over the course of almost two years, the court weighed motions and responses regarding which law should be applied to which issues. In ruling upon these motions, the court first had to determine what state's choice-of-law rule would be used. The court found that different state's choice-of-law rules should be applied to different substantive choice-of-law analyses. For example, for those cases where the court's jurisdiction was based on the Foreign Sovereign Immunities Act, the court chose to apply a federal common law choice-of-law rule.[71] For those cases where the court's jurisdiction was based on diversity of citizenship, the court applied the choice-of-law rule of the forum state.[72] The majority of the cases' choice-of-law rules had adopted the Restatement and application of depecage.

At issue in *Roselawn III* was which substantive law regarding compensatory damages should apply to those passengers from Indiana who died in flight 4184. The defendants maintained that the law of Indiana should apply, while the plaintiffs urged that either Illinois or Texas law should apply.[73] Definitive differences in state law existed between Indiana and Illinois/Texas law, such as whether siblings of decedents were allowed to recover damages and whether a plaintiff could have brought survival claims for conscious pain and suffering.[74]

---

67. 926 F. Supp. 736, 740 (N.D. Ill. 1996) [hereinafter Roselawn III].

68. *Id.* at 737-38.

69. *Id.*

70. *In re* Aircrash Disaster Near Roselawn, Indiana on October 31, 1994, No. 95 C 4593, MDL 1070, 1997 WL 572897, at *2-4 (N.D. Ill. Sept. 9, 1997) [hereinafter Roselawn V].

71. *Roselawn III*, 926 F. Supp. at 739.

72. *In re* Air Crash Near Roselawn, Indiana on October 31, 1994, 948 F. Supp. 747, 753-54 (N.D. Ill. 1996) [hereinafter Roselawn IV].

73. *Roselawn III*, 926 F. Supp. at 741.

74. *Id.* at 741-42.

In its application of depecage to this issue, the court went straight down the list of factors set forth in the Restatement (Second) of Conflicts section 145(2): (1) the place where the injury occurred; (2) the place where the conduct causing the injury occurred; (3) the domicile, residence, nationality, place of incorporation and place of business of the parties; and (4) the place where the relationship, if any, between the parties is centered.[75] As to factor one, the court obviously found that Indiana was the place where the injury occurred.[76] Concerning factor two, the court found that with regard to compensatory damages, the place where the conduct causing the injury occurred has little significance.[77] However, the court noted plaintiffs' claims that several acts of negligence occurred at American Airlines headquarters in Texas.[78] Factor three was easily decided by the court because all of the decedents were residents of Indiana with all of their estates pending in Indiana as well.[79] The court's analysis of factor four found that there was no real relationship established between the decedents and ATR or American Eagle.[80] Most of the decedents' tickets were purchased through a travel agent, and suggested if any relationship existed it revolved around the place of departure, Indianapolis, Indiana.[81] Therefore, the court found that Indiana had the greatest interest in determining and providing for the appropriate recovery for survivors and estates of Indiana decedents.[82]

The court found the domicile of the decedents to be the greatest factor in determining which state's law to apply. The court discussed the rationale for giving this factor more weight when considering the issue of compensatory damages:

> The legitimate interests of [the domiciliary states], after all, are limited to assuring that the plaintiffs are adequately compensated for their injuries and that the proceeds of any award are distributed to the appropriate beneficiaries. . . . Those interests are fully served by applying the law of the plaintiffs' domiciles as to issues involving the measure of *compensatory* damages (insofar as that law would enhance the plaintiffs' recovery) and the distribution of any award. Once the plaintiffs are made whole by recovery of the full measure of compensatory damages to which they are entitled under the law of their domiciles, the interests of those states are satisfied.[83]

Judge Castillo, in a separate opinion, similarly ruled that for those decedents

---

75. *Id.* at 742-44.

76. *Id.* at 742-43.

77. *Id.* at 743.

78. *Id.*

79. *Id.* at 743-44.

80. *Id.* at 744.

81. *Id.*

82. *Id.* at 744-45.

83. *Id.* at 745 (quoting *In re* Air Crash Disaster Near Chicago on May 25, 1979, 644 F.2d 594, 613 (7th Cir. 1981) [hereinafter Air Crash Chicago]).

who were non-Indiana residents, the substantive law relating to compensatory damages for pre-impact fear should be governed by each passengers' domicile. In so deciding the court reasoned that:

> [a]pplying the law of the injured person's domicile to issues of compensatory damages advances the principles of predictability and the protection of justified expectations set out in Second Restatement § 6, in that it respects the decedent's deliberate choice to make his or her home in a state and be governed by the laws of that state. Especially with respect to the claims of a decedent's estate, which are traditionally governed by the laws of the decedent's domicile, this deliberate decision to submit the daily affairs of life to the laws of a particular state may create justifiable expectations worthy of protection.[84]

Castillo further dissected the Roselawn litigation by separating out the issue of punitive damages. The plaintiffs claimed that Texas law should be applied to their punitive damage claims against American Airlines and ATR, while the defendants sought the application of the laws of France and Indiana.[85] Again, a difference existed between the law of Texas, which permitted punitive damages in wrongful death and survival claims, and the laws of Indiana and France, which would not permit punitive damages.[86]

Under the same depecage analysis, the court applied the law of Texas to plaintiffs' punitive damages claims against American Airlines.[87] In so holding, the Court reasoned that "the state that is both the principal place of business *and* the site of much of the alleged misconduct, Texas, has a far stronger interest in seeing its punitive damages law applied than any one of the half-dozen states in which some isolated aspects of the alleged misconduct occurred."[88] As to ATR, the Court held that French law should apply to punitive damage claims for the exact same reasons that Texas law would apply to American Airlines.[89] ATR's principal place of business was in France, and all of the alleged misconduct relating to the design, testing and manufacture of the ATR 72 airplane occurred in France.[90]

Finally, as if *Roselawn* had not been sufficiently dissected, the court found that as to plaintiffs' punitive damage claims against ATR's domestic corporation located in Virginia, Texas law should apply.[91] A conflict existed between Texas' allowance of uncapped punitive damages and Virginia's cap on punitives at $350,000 per person.[92] Much like the true conflict pertaining to Northwest

---

84. *Roselawn IV*, 948 F. Supp. at 758.
85. Roselawn V, 1997 WL 572897, at *1 (N.D. Ill. Sept. 9, 1997).
86. *Id.*
87. *Id.* at *3.
88. *Id.*
89. *Id.* at *4
90. *Id.*
91. *Id.* at *5
92. *Id.*

Airlines which arose in *Detroit Metro*, a true conflict existed here because ATR's wrongdoing occurred in Texas, while its principal place of business was in Virginia.[93] The court looked to *Air Crash Chicago* for precedent on how to "break the tie" when the misconduct took place in one state (Texas) and the principal place of business is in another (Virginia).[94] *Air Crash Chicago* held that when the interests of other jurisdictions are equal and opposed, the law of the place of injury should be applied.[95]

However, the court did not use this "tie breaker" rule because it's application would result in the use of Indiana law as the place of injury.[96] As Indiana law would not permit punitive damages, this would only serve to frustrate and undermine the policies behind both Virginia and Texas law.[97] To apply this sort of tie breaking test would put courts back into the arbitrary analysis of *lex loci delicti*. The court looked to comments found in the Restatement (Second) of Conflicts section 145 to determine that the law of Texas should apply because that is where the misconduct of ATR's domestic corporation took place.[98]

The case presented in *Schoeberle v. United States* involved an aircrash that killed a pilot and his two passengers.[99] All plaintiffs asserted claims against the United States of America under the Federal Tort Claims Act (FTCA) and against Signature Flight Corporation. The plaintiffs representing the deceased passengers also brought a claim against Monarch Aviation Services, the company who owned the aircraft, and the estate of the deceased pilot.[100] The plane crashed en route from Cedar Rapids, Iowa, to Milwaukee, Wisconsin, near a farm in Bernard, Iowa. The pilot had been having problems with the plane en route to Cedar Rapids and decided to have Signature Flight perform maintenance on the aircraft before returning to Milwaukee.

However, shortly after departing from Cedar Rapids, the plane began to experience severe engine problems, and the pilot contacted air traffic controllers (United States employees) located in Chicago.[101] Although there was an airport only 6.5 miles from the plane's position, controllers advised the pilot not to proceed to this airport because it did not have an instrument flight approach which would enable the pilot to guide his plane to the ground safely through the low clouds and poor visibility.[102] The plane traveled another 7.5 miles and then crashed, never reaching the controller's recommended airport in Dubuque, Iowa. It was later discovered that at the time of the accident, the airport which air

---

93. *Id.*; *see supra* notes 62-66 and accompanying text.

94. *Id.* at *5.

95. *See* Air Crash Chicago, 644 F.2d 594, 615-16 (7th Cir. 1981).

96. *Roselawn V*, 1997 WL 572897 at *5.

97. *Id.*

98. *Id.*

99. Schoeberle v. United States, Nos. 99 C 0352, 99 C 2599, 99 C 2292, 2000 WL 1868130 at *1 (N.D. Ill. Dec. 18, 2000).

100. *Id.*

101. *Id.* at *1-2.

102. *Id.* at *2.

traffic controllers thought was without an instrument approach did in fact have an instrument approach, which would have enabled the pilot to guide the plane safely to the ground.[103]

In analyzing what law to apply, the *Schoeberle* court was required under the FTCA to use the law of the state where the government's negligent act or omission occurred, which in this case was Illinois, the location of the air traffic controllers.[104] In applying Illinois choice-of-law principles, the court used depecage to wade through the complex but interesting connection the plane and its passengers had with three different states: Illinois, Wisconsin, and Iowa.[105] In its analysis, the court first examined what law would apply to issues of defendants' liability. Relying upon the Restatement (Second) of Conflicts section 175, the court applied the presumption that the law of the place of injury governs issues of liability in a wrongful death action, and therefore ruled that Iowa law must apply to the liability of Monarch Corp., Signature Flight Corp., and the pilot.[106] However, as to the United States' liability for the negligence of its air traffic controllers, the Court ruled that Illinois law must govern.[107]

In so holding, the court looked to the Restatement (Second) of Conflicts,[108] which defines the place of injury as "the place where the force set in motion by the actor first takes effect on the person. This place is not necessarily that where the death occurs."[109] Although the pilot's misconduct took place in Iowa where the aircraft crashed, this choice-of-law question involved the alleged misconduct of the United States which occurred in Illinois.[110]

The court determined the applicable law for compensatory and punitive damages separate from liability. The *Schoeberle* court followed *Roselawn III & IV* in finding the domicile of the decedents (Wisconsin) should govern compensatory damage claims.[111] As to plaintiffs' punitive damage claims against Signature Flight Corp., Monarch Corp., and the pilot, the court found that Iowa law should apply.[112] Again, similar to *Roselawn V*, the court considered both the place of injury and the place where the misconduct occurred as the rationale for their decision.[113]

---

103. *Id.*
104. *See* Richards v. United States, 369 U.S. 1, 6-11 (1962).
105. *Schoeberle*, 2000 WL 1868130 at *2-3.
106. *Id.* at *4-6.
107. *Id.* at *9.
108. RESTATEMENT (SECOND) OF CONFLICTS § 17 cmt. b (1971).
109. *Schoeberle*, 2000 WL 1868130 at *8.
110. *Id.*
111. *Id.* at *10-11.
112. *Id.* at *13-14.
113. *Id.*

C. Summary of the Choice-of-Law Analysis Used Under the Principles of Depecage

Although those cases which involve in-depth choice-of-law issues are often complex, there is a pattern to the analysis that has developed. Courts evaluate which choice-of-law rule they should use, whether there is a conflict among state laws, and then decide what state's substantive law should apply to each issue.

1. Whose Choice-of-Law Rule Should the Court Use?—This question ultimately asks whether the principles of depecage may be applied in the first place. If the court uses a state's choice-of-law rule that has not rejected *lex loci delicti*, then depecage may not be used.[114] However, as later discussed, some states have adopted modified forms of *lex loci delicti*, which further complicates the question as to whether depecage may be applied.[115] Are there any federal statutes that may dictate what state's choice of law must be used? For example, under the Federal Tort Claims Act the whole law of the state where the "act or omission occurred," including its choice-of-law rule, must be used.[116] If the federal court has jurisdiction based on diversity, then the court must use the choice-of-law rule of the forum state.[117]

2. Is There a Conflict in the Laws?—For the analysis to continue, the states which may have an interest in seeing their law applied to a particular issue must have conflicting laws. A "true conflict" is one where the states with interests in the question presented have different laws, and the application of one state's law would conflict with or impair the interests of another state.[118] "A false conflict exists if only one jurisdiction's governmental interests would be impaired by the application of the other jurisdiction's law. In such a situation, the court must apply the law of the state whose interests would be harmed if its law were not applied."[119] However, as *Schoeberle, Roselawn III, IV, & V*, and *Detroit Metro* demonstrate, choice-of-law rules do not distinguish between true conflicts and false conflicts.[120] Courts continue to analyze "false conflicts" when most legal scholars agree that they involve no choice-of-law problem at all.[121]

3. What State's Substantive Law Should Apply to a Particular Issue?—In this final step courts must examine the Restatement section 145 factors: (1) the

---

114. *See* Roselawn IV, 948 F. Supp. 747, 754-55 (N.D. Ill. 1996).

115. *See* Schalliol v. Fare, 206 F .Supp. 2d. 689 (E.D. Pa. 2002); Allen v. Great Am. Reserve Ins. Co., 766 N.E.2d 1157 (Ind. 2002); Northwest Pipe Co. v. Eighth Judicial Dist. Ct., 42 P.3d 244, 245-46 (Nev. 2002); Motenko v. MGM Dist. Inc., 921 P.2d 933, 935-36 (Nev. 1996).

116. 28 U.S.C. § 1346(b)(1) (2000); Richards v. United States, 369 U.S. 1, 6-11 (1962).

117. *See* Erie R.R. Co. v. Tompkins, 304 U.S. 64, 78 (1938).

118. *See* Air Crash Chicago, 644 F.2d 594, 615 (7th Cir. 1981) (reexamining the apparent conflict of laws reveals no way in which conflict can be resolved by a restrained or moderate interpretation, conflict is indeed a "true" conflict).

119. *See* Lacey v. Cessna Aircraft Co., 932 F.2d 170, 187 (3d Cir. 1991).

120. Christian L. Wilde, *Dépecage in the Choice of Tort Law*, 41 S. CAL. L. REV. 329, 342-45 (1968).

121. *Id.*

place where the injury occurred; (2) the place where the conduct causing the injury occurred; (3) the domicile of the parties (in the case of corporations, this would include the place of incorporation and principal place of business); and (4) the place where the relationship between the parties is centered, if that can be determined.[122] To determine which factors to weigh more heavily, courts look to case law on a particular issue. For example, in determining what state's law will apply to compensatory damages, courts consider the domicile of the decedent or injured plaintiff to be paramount.[123]

As seen by these examples of courts applying depecage, there are definite benefits realized through its use. Depecage reveals and disposes of false conflicts where the contacts and/or interests of one state greatly outweigh another state's contacts or interests on a separate issue. Although bodies of law may differ, a false conflict does not present a choice-of-law problem because one state's law may be applied without undermining the other disinterested state's laws or policies.[124] Thus, a court can focus on the true conflicts where multiple states have a significant interest in seeing their law applied to a particular issue and where the application of one state's law will undermine the laws and policy of other interested states.

III. RATIONALE FOR REJECTING DEPECAGE

Without delving into the alleged flaws contained behind the theory of depecage, it is easy to see why some courts are hesitant to trod down the path of depecage. First, depecage forces the court's legal analysis to become extremely complex. Secondly, there is a question as to how far a court should split issues. Splitting liability from damages, punitive damages from compensatory damages, and then further dissecting these issues as they relate to individual parties may water-down the decisive perception that courts should possess. This type of issue splitting presents the picture of Russian Matryoshka dolls where inside each doll is a smaller doll. Every time a court examines an issue to determine which state's law will apply to that issue, it then finds an issue inside of that issue.

Most recently, "a counterrevolution of sorts appears to be emerging, marked by the insistence that the concept of rights should have a greater role to play."[125] Although the United States Supreme Court has stated otherwise, many scholars believe that the Constitution has something to say about uniformity in choice-of-law analysis.[126] These scholars argue that the Privileges and Immunities Clause[127] "destroys the domiciliary-centered conception of governmental

---

122. RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 145 (1971).

123. *See* Roselawn III, 926 F. Supp. 736, 745-46 (N.D. Ill. 1996).

124. *See* Wilde, *supra* note 120.

125. Kermit Roosevelt III, *The Myth of Choice of Law: Rethinking Conflicts*, 97 MICH. L. REV. 2448, 2466 (1999).

126. *Id.* at 2466-67.

127. U.S. CONST. art. IV, § 2.

interests."[128] If a state grants rights to its domiciliaries, it must grant them to nondomiciliaries in the same cases. "The Privileges and Immunities Clause thus prevents the crudely selective exercise of legislative jurisdiction to favor domiciliaries."[129] This constitutional approach would eliminate the need for depecage in its current form.

Included in the push to alter the concept of depecage is an acknowledgment of the danger of distorting legislative intent by choosing certain portions of a state's laws to apply to a particular issue.[130] "[L]egislators may enact a given law only because of its expected interaction with a complementary law. [It would be] inappropriate to apply a state's wrongful death rule without its damage cap, which may have been an important condition on the adoption of the wrongful death statute."[131] For example, state A might provide broad recovery for injuries but adopt defenses or immunities to prevent fraudulent claims, which prevents too broad a recovery. In contrast, state B might permit all injured plaintiffs to sue for compensation, but prohibit direct actions, cap damages, limit negligence per se, and apply a narrow res ipsa loquitur doctrine.[132]

These bundles of law reflect different legislative decisions providing the optimal combination of legal standards deemed appropriate for that state. Plaintiffs should not be allowed to put "together half a donkey and half a camel, and then ride to victory on the synthetic hybrid."[133] How can, on the one hand, depecage emphasize governmental interests and, on the other, accept an outcome that none of the supposedly interested states would condone.[134] Additionally, when laws remain bundled together, any party need only make a single prediction as to what law may apply rather than making the separate, multiple predictions that depecage requires.[135] In order for people to abide by the law, they have to know which law applies. Separation of issues under depecage makes it much more difficult to predict which law a court will apply to each issue for any alleged misconduct.

In some ways depecage is no better than *lex loci delicti*. Like the *lex loci delicti* vested rights theory, depecage interest analysis avoids the difficult task of resolving conflicts between laws, though in a somewhat different way. *Lex loci delicti* denies the possibility of conflict, allowing only one law to govern the transaction. Depecage admits the conflict. Indeed, depecage recognizes the

---

128. Roosevelt, *supra* note 125, at 2534.

129. *Id.*

130. *See* William H. Allen & Erin A. O'Hara, *Generation Law and Economic of Conflicts of Laws: Baxter's Comparative Impairment and Beyond*, 51 STAN. L. REV. 1011, 1034-36 (1999); Erin A. O'Hara & Larry E. Ribstein, *From Politics to Efficiency in Choice of Law*, 67 U. CHI. L. REV. 1151, 1192-93 (2000).

131. O'Hara & Ribstein, *supra* note 130, at 1193.

132. *Id.*; Allen & O'Hara, *supra* note 130, at 1034-36.

133. Frederick K. Juenger, *How Do You Rate a Century?*, 37 WILLIAMETTE L. REV. 89, 106 (2001) (quoting Brainerd Currie).

134. *Id.*

135. O'Hara & Ribstein, *supra* note 130, at 1193-94.

distinction between true conflict and false conflict. However, even depecage reverts to an arbitrary analysis for true conflicts where multiple states have equal interests in the application of their law. It is here where interest analysis under depecage fails by employing a technique that suggests that these true conflicts do not need to be resolved.[136] Reverting to an altered form of *lex loci delicti* analysis may eliminate the need for depecage. For example, *lex loci delicti* "could be modified to include a common domicile rule for 'loss distribution' rules while retaining a 'place of injury' rule for those laws directed toward conduct."[137]

Depecage has been labeled as an escape device to often harsh choice-of-law determinations.[138] "Because formal rules can never capture the complexities or depth of [every] situation they are meant to govern, they have a tendency to multiply exponentially," thus producing escape devices.[139] Escape devices, such as depecage, while producing just results, are intellectually dishonest. "They . . . tend to produce narrow rules that are not helpful in other cases with different facts."[140] Thus, depecage as it coexists with interest analysis, may give courts too much discretion. This is especially true in single-accident mass-tort cases, where the interests of multiple states are at issue. Depecage only compounds the wide discretion courts have under a state interest analysis, and thereby allows courts more discretion over their resolution of choice-of-law questions.[141]

Depecage also threatens neutrality and equality. Neutrality could be threatened through a court's abuse of depecage combining states' laws on various issues to create a set of rules that is more favorable to one party than the law of any single state.[142] Equality among parties is also in danger.

> [D]epecage favors the party with access to greater legal resources (usually defendants) because it requires a separate state-interest analysis as to each legal issue in the case, which must be multiplied by the number of interested states. It is thus much simpler, yet no less fair in single-accident mass-tort cases, to select one state's law to govern all issues in the entire controversy.[143]

One court has even compared the use of depecage to a type of lottery system.[144] The "winners" of the lottery would be those injured by tortfeasors

---

136. Roosevelt, *supra* note 125, at 2463.

137. Allen & O'Hara, *supra* note 130, at 1044.

138. Erich H. Gaston, *Reassessing Connecticut's Eclectic Choice of Law Methodology: Time for (Another) New Direction*, 73 CONN. B.J. 462, 465-66 (1999).

139. *Id.* at 466.

140. *Id.*

141. Thomas M. Reavley & Jerome W. Wesevich, *An Old Rule for New Reasons: Place of Injury as a Federal Solution to Choice of Law in Single-Accident Mass-Tort Cases*, 71 TEX. L. REV. 1, 38 n.184 (1992).

142. *Id.*

143. *Id.*

144. *See* Boomsma v. Star Transport, Inc., 202 F. Supp. 2d 869, 879 (E.D. Wis. 2002).

from a foreign state that has liberal liability and damage rules, such as no cap on wrongful death damages. The "losers" would be those injured by fellow citizens of their own conservative state where recovery is limited. Such a system would undermine certainty, predictability, and uniformity of result.[145]

IV. LACK OF DIRECTION FOR THE DOCTRINE OF DEPECAGE

Despite the recent scholarly push to eliminate the use of depecage, courts have not followed suit. However, while courts have not affirmed the elimination of depecage, neither have they provided any direction for the doctrine. With the exception of air crash disaster litigation, very few courts have provided access for application of the doctrine. Thus, one reason for the courts' failure to embrace the use of depecage in choice-of-law analyses is the lack of precedent available from higher authority.

There are few written legal opinions from appellate and supreme courts which concern the doctrine of depecage. In fact, the U.S. Supreme Court has never used the term depecage in a written opinion.[146] Conflicts as a whole has been abandoned by the Supreme Court flowing from the recognition of the difficult and complex nature of conflict of law issues, and from the Court's current views to err on the side of federalism rather than nationalism.[147] Almost as lacking are written opinions from the circuit court of appeals. Only twenty-three circuit court of appeals written opinions exist which have used the term depecage, while there are literally hundreds of opinions written by the court of appeals which discuss choice of law.[148] State supreme courts have not fared much better in explaining the doctrine of depecage.[149]

Unfortunately, some courts have used a form of depecage without expressly stating so, or explaining the reasoning and source behind their analysis. In *Allen v. Great American Reserve Insurance Co.*, the Indiana Supreme Court applied a form of depecage without expressly stating so.[150] Thus, when courts look to Indiana's choice-of-law rule, they are left in a state of uncertainty. Choice-of-law analysis is difficult enough without courts implementing legal rules which lead to a lack of clarity and continuity. In *Allen*, North and South Carolina insurance agents brought an action against the life insurer, Great American Reserve Insurance Co. ("GARCO") and the general agent, Glen Guffey, for

---

145. *See id.*

146. Search performed on February 25, 2003 of WestLaw databases "SCT" and "SCT-OLD" using search term "depecage."

147. Roosevelt, *supra* note 125, at 2503-04.

148. Search performed on February 25, 2003 of Westlaw database "CTA" and "CTA-OLD" using search term "depecage."

149. *See* Allen v. Great American Reserve Ins. Co., 766 N.E.2d 1157 (Ind. 2002); *see also* Brown v. Kleen Kut Mfg. Co., 714 P.2d 942, 943-46 (Kan. 1986) (claiming to adhere to *lex loci delicti* while separating out issues of successor corporation liability and tort law).

150. *Allen*, 766 N.E.2d at 1157.

misrepresentations about front-end load annuities.[151] Twelve counts were asserted against the two defendants alleging claims for breach of contract, fraud, Indiana Crime Victims Relief Act, Indiana statutory fraud, Indiana statutory deception, Indiana criminal mischief, North Carolina unfair trade practices, South Carolina unfair trade practices, civil conspiracy, tortious interference with a business relationship, negligence, indemnification, and accounting.[152] The Indiana Supreme Court reversed the trial court's application of South Carolina law to all of the claims.[153] In so ruling, the *Allen* court held that

> because the parties hail from different states, and because many of the activities in question occurred in different states, this case raises significant choice of law issues. In analyzing each of the counts of the plaintiffs' complaint, it is first necessary to determine which state's law applies to that count. The answer may differ for different counts and may differ between defendants as to a single count.[154]

In applying this rule, the court used various state's laws to different counts, and used the Restatement factors to analyze the contacts of the various states.[155] For example, plaintiffs' count I claim against GARCO for breach of a covenant of good faith implied in their contracts with GARCO was controlled by Indiana law.[156] Conversely, plaintiffs' count I claim against Guffey for a breach of the duty of good faith and fair dealing arising from the agency relationship itself, apart from any claims of breach of contract, was governed by South Carolina law.[157]

At first glance, it appears clear that *Allen* embraces the principles of depecage. However, *Allen* never uses the term "depecage," nor does it state that it is applying different state's law to different issues. *Allen* only purported to apply different state's law to different counts and different parties, and did not go so far as to separate out issues of damages and liability.[158] What then is

---

151. *Id.* at 1160-61.

152. *Id.* at 1162-70.

153. *Id.* at 1161-62, 1170.

154. *Id.* at 1162.

155. RESTATEMENT (SECOND) OF CONFLICTS § 145 (1971).

156. *Allen*, 766 N.E.2d at 1162.

157. *Id.* at 1163.

158. Simon v. United States, 341 F.3d 193, 201 (3d Cir. 2003) ("*Allen* was a routine application of different choice-of-law analyses to different *counts*, as opposed to different issues within a single count."). Indiana may be forced to clarify its choice-of-law rule as *Simon* has certified two questions, which have arisen out of *Schalliol*, to the Indiana Supreme Court. First, the Indiana Supreme Court is to determine if a true conflict exists between Indiana's and the District of Columbia's choice-of-law rules. *Id.* at 205. In other words, the court must decide if Indiana recognizes depecage. *Id.* at 195-96. Secondly, the court must decide "how to resolve a split among the *Hubbard* factors in choosing a jurisdiction's substantive law when one factor points toward Indiana, another toward Pennsylvania, and the third is indeterminate . . . ." *Id.* at 205. The Indiana Supreme Court's rulings on these certified questions could significantly reshape Indiana's choice-

Indiana's stance on depecage? Part of the problem lies with Indiana's partial adoption of the Restatement (Second) of Conflicts. Indiana has never expressly adopted the Restatement as a whole but has established a modified *lex loci delicti* standard that looks first to the place of the tort to determine if there is any significant connection to the legal action.[159] However, in Indiana's defining choice-of-law case, *Hubbard Manufacturing Co. v. Greeson*, the court did adopt the Restatement section 145, which specifically states that "[t]he rights and liabilities of the parties with respect to an issue in tort are determined by the local law of the state . . . with respect to that issue . . . ."[160] Indiana is left in a state of uncertainty as to what its choice-of-law rule really is.

Indiana is not alone in failing to clearly set forth a framework for the use of depecage. Nevada, which is also a "twilight zone" state somewhere between *lex loci delicti* and the Restatement (Second) of Conflicts analysis, has also failed to grasp an opportunity to give guidance to lower courts concerning the doctrine of depecage. In *Northwest Pipe Co. v. Eighth Judicial District Court*, the court failed to embrace the concept of depecage.[161] The underlying actions in *Northwest Pipe* arose from an accident that occurred on a highway in San Bernardino County, California, when three concrete pipes weighing several tons fell off a Northwest Pipe Company truck and struck several vehicles. Six individuals were killed: two Nevada residents, and four California residents.[162] Northwest Pipe Company was an Oregon corporation doing business throughout the United States.[163] Suit was brought in Nevada, and Nevada's choice-of-law rule found in *Motenko v. MGM Distributor Inc.* was used.[164] *Motenko* held that the law of the forum should apply unless two or more of the four factors enumerated in the Restatement section 145 show that another state has an overwhelming interest in the litigation.[165] The court in *Northwest Pipe* was deeply divided, with two justices concurring in part and dissenting in part, and two justices dissenting.[166] The majority opinion held that Nevada law should be applied to all plaintiffs; however, the dissenters comprised a majority as to the California plaintiffs and held that California law should be applied to those cases.[167] Surprisingly, none of the justices in *Northwest Pipe* even mentioned the possibility of applying depecage; yet in a sense this is what the court did in applying the law of California to the California plaintiffs. True depecage in this case would have looked something like this: 1) as to issues of liability, California

---

of-law rule.

159. Hubbard Mfg. Co. v. Greeson, 515 N.E.2d 1071, 1074 (Ind. 1987).

160. *Id.*; RESTATEMENT (SECOND) OF CONFLICT OF LAWS: THE GENERAL PRINCIPLE § 145 (1971).

161. *See* Northwest Pipe Co. v. Eighth Judicial Dist. Court, 42 P.3d 244, 245-46 (Nev. 2002).

162. *Id.* at 245.

163. *Id.*

164. *Id.*; Motenko v. MGM Distrib. Inc., 921 P.2d 933 (Nev. 1996).

165. *Motenko*, 921 P.2d at 935.

166. *Northwest Pipe Co.*, 42 P.3d at 246, 249.

167. *Id.* at 246, 248.

or Oregon law would apply; 2) as to issues of compensatory damages for the Nevada residents, Nevada law would apply; 3) as to issues of compensatory damages for the California residents, California law would apply. Although *Northwest Pipe* does not go into details about the possible conflicts existing between California and Nevada law, simply from a state interest stand point, depecage is a logical solution to the court's divided dilemma.

The Seventh Circuit recently ignored an opportunity to apply the doctrine of depecage in *In re Bridgestone/Firestone, Inc.*[168] In *In re Bridgestone/Firestone*, the court was concerned with the issue of whether a class action certified by the District Court for the Southern District of Indiana met the commonality and superiority requirements of Federal Rules of Civil Procedure 23(a), (b)(3).[169] As a part of its analysis, the court had to struggle with the choice-of-law issue because no class action is proper unless all litigants are governed by the same legal rules.[170] Using Indiana's choice-of-law rule, the district court held that Indiana would look to the headquarters of the defendants, because that is where the products were designed and where the important decisions about disclosures and sales were made.[171] This ruling meant that all claims by the Ford Explorer class would be resolved under Michigan law (Ford Headquarters) and all claims by the tire class would be resolved under Tennessee law (Firestone Headquarters).[172] Thus, appropriate uniform law would serve to satisfy the commonality requirement.

Nevertheless, the Seventh Circuit reversed the district court's ruling and held that a certifiable class did not exist because the proper choice-of-law analysis would require the court to apply fifty different state laws to the issues presented.[173] Thus, no commonality existed.[174] In so holding, the court stated that "[i]t follows that Indiana's choice-of-law rule selects the [fifty] states and multiple territories where the buyers live, and not the place of the sellers' headquarters, for these suits."[175] The court hinted that there may be a better way to approach the problem presented in *In re Bridgestone/Firestone* but failed to elaborate upon their thinking.[176] The court concluded that they were bound by what they interpreted as Indiana's *lex loci delicti* rule.[177] Put in simple terms, the Seventh Circuit failed to recognize that Indiana's choice-of-law rule, specifically *Allen v. Great American Reserve Insurance, Co.*, may well call for the use of depecage.

---

168. *In re* Bridgestone/Firestone, Inc., 288 F.3d 1012 (7th Cir. 2002), *cert. denied*, 123 S. Ct. 870 (2003).

169. *Id.* at 1015.

170. *Id.*

171. *Id.*

172. *Id.*

173. *Id.* at 1016, 1018.

174. *Id.*

175. *Id.* at 1018.

176. *Id.* at 1019-20.

177. *Id.* at 1016.

In so doing, the Seventh Circuit missed a perfect opportunity to promote efficient litigation, while still respecting state sovereignty. If the *In re Bridgestone/Firestone* court was truly interested in fairly compensating the injured plaintiffs, then a proper result could have been reached by applying the law of Tennessee to issues of liability against Firestone, the law of Michigan to issues of liability against Ford, and the law of each injured parties' domicile for determining the measure of damages.

This is precisely the approach the district court took in *Simon v. Philip Morris Inc.*[178] In *Simon*, the plaintiffs' core theory sounded in fraudulent concealment by the tobacco industry, and consisted of a nationwide class implicating the interests of all fifty states' laws. Some of those laws conflicted with New York's laws (Philip Morris Headquarters).[179] However, rather than applying the law of the state where the plaintiff resided or had purchased the majority of his or her cigarettes, the court applied the law of New York to issues of liability and punitive damages.[180] The court reasoned that a number of critical meetings of tobacco representatives necessary to orchestrate the scheme allegedly occurred in New York, and at least two of the companies, Lorillard and Philip Morris, Inc., had their principle places of business in New York.[181] However, the court did not end its analysis there; it applied depecage to separate issues of liability and damages. In so deciding the court stated that:

> In this action, state interests of states other than New York will be less implicated in any conflict since the court envisions transferring individual compensatory questions to each plaintiff's home district. Through the use of depecage, each claimant will rely upon his or her own state law with regard to critical individual recovery issues.[182]

The court went on to add that "[w]hile New York has a paramount interest in punishing and deterring misconduct in New York, other states have a concurrent interest in ensuring that their own citizens receive individual relief in line with their own compensatory schemes."[183]

In re Bridgestone/Firestone is analogous to *Simon* in many respects. Both cases proposed theories of fraud, both involved a single principal place of business where most of the fraudulent conduct occurred, and both involved the purchase of a nationwide product. In *Simon*, as the place where a plaintiff bought his or her cigarettes bore little significance to the litigation, so too does the place where a plaintiff bought his or her tire or vehicle control all issues in

---

178. Simon v. Philip Morris, Inc., 124 F. Supp. 2d 46, 71-77 (E.D.N.Y. 2000); Elizabeth J. Cabraser, *Unfinished Business: Reaching the Due Process Limits of Punitive Damages in Tobacco Litigation Through Unitary Classwide Adjudication*, 36 WAKE FOREST L. REV. 979, 1012-13 (2001).

179. *Simon*, 124 F. Supp. 2d at 71.

180. *Id.* at 72.

181. *Id.*

182. *Id.* at 74.

183. *Id.* at 76.

Bridgestone/Firestone. The only possible difference between these two situations which may prevent a court from applying depecage is that in *In re Bridgestone/Firestone*, the plaintiffs were trying to avoid the return of their individual cases to the originating court.[184] If, after a determination on liability or punitive damages is reached as a certified class, the individual cases are not returned for the calculation of damages under the law of each plaintiffs' domicile, then splitting issues of liability and compensatory damages would be pointless.

With these examples of courts fumbling with the doctrine of depecage, a clear analysis of the rationale supporting depecage is needed. Those courts that are teetering on the edge of modern conflicts thinking should openly embrace depecage for the following reasons.

V. DEPECAGE: A NEEDED TOOL FOR EFFECTIVE CHOICE-OF-LAW ANALYSIS

The rationale for the "center of gravity" or "grouping of contacts" doctrine adopted by the *Babcock* court, accommodating competing state interests in tort cases with multi-state contacts, centers on justice, fairness, and achieving the best practical result. The best practical result is reached "by giving controlling effect to the law of the jurisdiction which, because of its relationship or contact with the occurrence or the parties, has the greatest concern with the specific issue raised in the litigation."[185]

A. A Logical Basis for Depecage

The merit behind the separation of issues is that it gives to the place having the most interest in the problem paramount control over the legal issues arising out of a particular factual context, and thereby allows the forum to apply the policy of the jurisdiction most intimately concerned with the outcome of the litigation.[186]

State rules of law are promulgated for a reason. They embody social, economic, or administrative policies which can be discovered through the ordinary way courts interpret and construct laws.[187] The rule of vested rights and *lex loci delicti* are void of this process for which courts were designed. Under the archaic vested rights doctrine, courts apply a law whose underlying policy would not be furthered by its application. In truly perverse fashion, this doctrine could even operate to suppress the interest of one state without advancing the interests of the other.[188] The doctrine of depecage, on the other hand, purports that because every rule of law may have a different purpose they must be construed separately. Depecage breaks down choice-of-law problems into

---

184. *See In re* Bridgestone/Firestone, Inc., 288 F.3d 1012, 1015 (7th Cir. 2002).

185. *Simon*, 124 F. Supp. 2d at 55; Babcock v. Jackson, 191 N.E.2d 279, 281-82 (N.Y. 1963).

186. *Babcock*, 191 N.E.2d at 283.

187. Southerland, *supra* note 23, at 479.

188. *Id.*

smaller groups to facilitate more adequate analysis of underlying policies.[189]

Depecage provides the means to solve complex choice-of-law problems which may be impossible or at least extremely difficult to solve under a traditional choice-of-law framework. Splitting the choice-of-law process into as many choices of law as there are conflicting rules has the advantage of reducing the number of seemingly unsolvable conflicts which exist between two bodies of state law taken as a whole.[190] Depecage seeks to apply to each issue the rule of the state with the greatest concern in the determination of that issue. Depecage also serves to effectuate the purpose of each of the rules applied without disappointing party expectations.[191]

B. Flaws in the Counter Movement's Approach to Choice-of-Law and Depecage

The counterrevolution to modern choice of law, which seeks to return to a form of vested rights analysis, has failed to see the importance of differing state laws and how depecage minimizes true conflicts.[192] Those who still believe in vested rights and *lex loci* have stated that the Supreme Court should intervene, and still others have gone a step farther stating that Congress should enact arbitrary rules on how to decide state conflicts of law issues.[193]

The first flaw in this reasoning is that it seeks uniformity and predictability in choice-of-law analysis at any cost; the cost of which is too high.[194] Unpredictability leads to real benefits from diversity and innovation among state law.[195] "[S]tate innovation is no judicial myth. When Wyoming became a State in 1890, it was the only State permitting women to vote. That novel idea did not bear national fruit for another thirty years. Wisconsin pioneered unemployment insurance, while Massachusetts initiated minimum wage laws for women and minors."[196] Americans and our legal system have learned to live with a certain amount of unpredictability in the application of law. Uniformity through federal legislation, pertaining to choice of law, would cost citizens the ability to govern, allowing them only to administer their local problems.[197]

---

189. Wilde, *supra* note 120, at 345.

190. *Id.* at 346.

191. Willis L. M. Reese, *Depecage: A Common Phenomenon in Choice of Law*, 73 COLUM. L. REV. 58, 60 (1973).

192. Symposium, *Choice of Law: How It Ought to Be, A Roundtable Discussion*, 48 MERCER L. REV. 639, 705 (Winter 1996) (speaking against congress becoming involved in state choice-of-law analysis).

193. *Id.* at 703-04.

194. *Id.* at 705; *see* Roosevelt, *supra* note 125; *see also* Reavley & Wesevich, *supra* note 141, at 2-8.

195. Symposium, *supra* note 192, at 705.

196. Fed. Energy Regulatory Comm'n v. Mississippi, 456 U.S. 742, 788-89 (1982) (O'Connor, J., concurring in part and dissenting in part) (footnotes omitted).

197. *Id.* at 789.

Secondly, even those who want federal control over state choice-of-law analysis admit that such control should only be exerted where a true conflict exists.[198] Professor David Currie stated that he "would like to see Congress enact some rules, and . . . would be happy to have them be perfectly arbitrary—place of the wrong, if you will—but only once you conclude that there is a true conflict."[199] Some type of choice-of-law analysis must still be in place to determine what is a true conflict. Thus, the ability of depecage to isolate true conflicts between differing state's laws is still useful.

The reasoning that depecage distorts state legislative intent by picking and choosing only certain laws out of a broader area of law also contains certain flaws.[200] In *Detroit Metro*, defendant, McDonnell Douglas, argued that the alleged misconduct did not occur in California where the aircraft was partially manufactured and designed, because the three allegedly defective component parts at issue were respectively and individually designed and manufactured in Massachusetts, Missouri, and Arizona.[201] Thus, McDonnell Douglas urged the court to use three different state laws as to plaintiffs' design defect claims. The *Detroit Metro* court, however, stated that

> to pare the Plaintiffs' design defect claims into subcategories and thereby conduct a separate choice of law analysis for each component, which has been alleged to be defective, would complicate and obfuscate the complex choice of law issues in this case and would require the jury to apply various and, possibly different, product liability standards in the same case dependent upon the particular component at issue.[202]

Detroit Metro is an example of a court's recognition of where to draw the line on issue splitting. Inherent in the *Detroit Metro* court's self restraint is the notion that issue splitting can only go so far before the true meaning of a body of law is lost. As such, the court in *Detroit Metro* realized that applying three different product liability standards to the same case would have the potential to distort state interests.[203] Therefore, although there is a valid concern regarding depecage's ability to distort legislative intent, courts are able to draw the line in the sand when it comes to issue splitting, and thereby thwart any danger of contorting state interests.

Even if critics of depecage lack confidence in a court's ability to exercise self-restraint and to avoid distorting legislative intent, the use of depecage may be justified by choice-of-law rules that mandate such distortion.[204] "When a choice-of-law rule calls for application of a given rule [depecage], that rule

---

198. Symposium, *supra* note 192, at 704.

199. *Id.* Professor David Currie is the son of famed choice-of-law scholar Brainerd Currie. *See id.* at 639-40.

200. *See* Allen & O'Hara, *supra* note 130; O'Hara & Ribstein, *supra* note 130.

201. Detroit Metro, 750 F. Supp. 793, 799 n.12 (E.D. Mich. 1989).

202. *Id.*

203. *Id.*

204. Reese, *supra* note 191, at 74.

should usually be applied even at the risk of distortion. To do otherwise would deprive the rule of certainty of application, which is a vital attribute of a rule of law."[205]

Claims that depecage is merely an escape device that courts can use to favor one party over another also fall short.[206] When the Supreme Court has spoken on choice of law, it has clearly stated that there are certain restrictions prohibiting arbitrary and unfair decisions on applicable law.[207] Looking to the Constitution, the Court stated that "for a State's substantive law to be selected in a constitutionally permissible manner, that State must have a significant contact or significant aggregation of contacts, creating state interests, such that choice of its law is neither arbitrary nor fundamentally unfair."[208] Courts are not free to pick and choose state law for the benefit of one party over another without justifying such a result within the bounds of state contacts and interests. Thus, the danger of depecage being used as an arbitrary and inequitable tool by courts is no more likely to occur than any other constitutionally prohibited legal doctrine.[209]

Finally, depecage has been criticized for its inability to resolve true conflicts (where more than one state has an interest in seeing its law applied).[210] However, depecage does not purport to be a uniform solution to true conflicts of law. The value of depecage is its ability to enable courts to target their consideration of interests and policies more precisely.[211]

C. Depecage: A Solution for a Legal Impossibility

Depecage is essential for several reasons, but most importantly, there may be some unique situations where its use is imperative. *Schalliol v. Fare* provides the perfect example of the necessity of depecage by demonstrating a situation where failing to apply depecage leaves a court with a legal impossibility.[212] Further analysis of *Schalliol* may allow courts to rethink the importance of depecage in our legal system.

Schalliol involves a plane crash and the subsequent deaths of the pilot, John Fare, Sr., and three passengers. Dennis Schalliol and B. Kenin Hart were passengers on the plane, whose personal representatives of their estates brought

---

205. *Id.*

206. *See* Gaston, *supra* note 138.

207. Phillips Petroleum Co. v. Shutts, 472 U.S. 797, 818-19 (1985).

208. *Id.* at 818 (quoting Allstate Ins. Co. v. Hague, 449 U.S. 302, 312-13 (1981)).

209. The real threat of arbitrary and unfair application of state law exists in a single designated jurisdiction rule whereby state interests are disregarded and parties either are overcompensated or under compensated. *See* Robert A. Sedler & Aaron D. Twerski, *The Case Against All Encompassing Federal Mass Tort Litigation: Sacrifice Without Gain*, 73 MARQ. L. REV. 76, 87-90 (1989).

210. *See* Roosevelt 125, at 2463.

211. Peterson, *supra* note 7.

212. *See* Schalliol v. Fare, 206 F. Supp. 2d 689 (E.D. Pa. 2002); *see also* Simon v. United States, 341 F.3d 193 (3d Cir. 2003).

legal actions. The estates claimed negligence on the part of FAA(United States) air traffic controllers in Indianapolis, and FAA(United States) offices in Washington, D.C. Mary Schalliol, widow of Dennis and personal representative of his estate also brought negligence claims against the pilot John Fare, Sr., and the owner and operator of the aircraft, Hart Delaware Corporation.[213] As to those claims against the United States, the United States filed a motion for determination of choice of law claiming that under the Federal Tort Claims Act, 28 U.S.C. §§ 1346(b), 2671, the law of Indiana applied to all aspects of those actions brought by plaintiffs against the United States.[214] The Eastern District Court of Pennsylvania ruled in favor of the United States' motion stating "that Indiana substantive law applies to all claims pled against the United States under the FTCA."[215] However, the court certified its ruling for immediate appeal to the Third Circuit Court of Appeals pursuant to 28 U.S.C. § 1292(b).[216]

Schalliol presents numerous contacts with multiple states. While in flight, Pilot Fare contacted an air traffic controller at the Indianapolis Air Route Traffic Control Center, who cleared him for landing via the Simplified Directional Facility approach at Runway 4 (SDF 4) at the Somerset, Kentucky airport. Pilot Fare also possessed an Instrument Approach Procedure (IAP) for the SDF 4, which was published by the United States, in Washington, D.C., and contained information stating that the approach was in service. However, the navigational facility supporting the SDF approach was out of service indefinitely.[217] Due to a combination of these factors the pilot flew into a radio tower striking the aircraft's right wing and hitting the ground several hundred feet beyond the tower.[218] At the time of the accident, the domiciles of the majority of the *Schalliol* plaintiffs were Pennsylvania. Mary and Dennis Schalliol, and Ken Hart, were residents of Pennsylvania, while John Fare, Sr., whose son John Fare, Jr. filed a cross-claim against the United States, resided in New Jersey.[219]

In their various claims against the United States, plaintiffs contended that the publication of the IAP was negligent. All parties agreed that any such negligence occurred in Washington, D.C. Plaintiffs also alleged that the air traffic controllers were negligent, and all parties agreed that any such negligence of the controllers, including any failure to monitor the aircraft or to supervise personnel, occurred in Indiana.[220] Finally, all parties agreed that any negligence of Pilot Fare occurred in Ohio and Kentucky.[221]

One of the reasons *Schalliol* presents such a complex choice-of-law problem is because all parties agreed that, as to claims against Fare and Hart Delaware,

---

213. *Id.* at 691-93.
214. *Id.*
215. *Id.* at 691.
216. *Id.*
217. *Id.* at 691-92.
218. *Id.*
219. *Id.* at 700.
220. *Id.* at 692.
221. *Id.*

Pennsylvania law applies.[222] The reason being that Hart Delaware, the owner and operator of the plane, along with its sole employee, pilot Fare, have their principal place of business in Philadelphia, Pennsylvania.[223] This problem is further complicated by the fact that in claims against Hart Delaware and John Fare, Sr., a jury is the trier of fact, where as to claims against the United States, the judge is the trier of fact.[224]

With all of these multiple parties, contacts and interests, *Schalliol* appears to be a perfect candidate for applying depecage. Why then did the district court decline to apply the principles of depecage to *Schalliol*? The court, using Indiana's choice-of-law rule, agreed with the United States that Indiana courts, having never before explicitly applied depecage, would not apply depecage in the case at bar, and declined to infer a significant choice-of-law principle into Indiana law.[225]

Indiana evaluated the Restatement (Second) of Conflicts section 145 under *Hubbard Manufacturing Co. v. Greeson*.[226] *Hubbard* is the seminal choice-of-law case in Indiana. In *Hubbard*, Indiana adopted the factors enumerated in section 145, but arguably never specifically endorsed or applied the concept of depecage.[227] Nowhere in *Hubbard* does the court use the term depecage, nor did the court expressly separate out legal issues such as liability and damages.[228] *Hubbard* involved the job-related death of an employee in Illinois who was using a lift manufactured in Indiana.[229] Despite the fact that injury took place in Illinois, the court applied Indiana law. However, in so doing, the court applied Indiana law to damages as well as liability, even though Illinois was the site of the job-related death and of the job itself, as well as the provider of workmen's compensation to the survivors.[230]

The *Hubbard* test begins with the presumption that the substantive law of the place where the tort or place of injury occurred governs the case.[231] However, if the place of the tort bears little connection to the legal action, Indiana follows the most significant relationship test based on the Restatement section 145.[232] The factors for this significant relationship test include: 1) the place where the conduct causing the injury occurred; 2) the residence and place of incorporation

---

222. Appellant's Brief at 3, *Schalliol* (No. 02-3996).

223. *Schalliol*, 206 F. Supp. 2d at 701 n.34.

224. 28 U.S.C. § 2402 (2002) ("[A]ny action against the United States under section 1346 shall be tried by the court without a jury . . . .").

225. *Schalliol*, 206 F. Supp. 2d at 700. *See also* Simon v. United States, 341 F.3d 193, 205 (3d Cir. 2003) (certifying questions for the Indiana Supreme Court to answer concerning the clarification of Indiana's choice-of-law rule).

226. Hubbard Mfg. Co., Inc. v. Greeson, 515 N.E.2d 1071, 1073-74 (Ind. 1987).

227. *See generally id.* (failing to mention the possibilities for issue splitting).

228. *Id.*

229. *Id.*

230. *Id.*

231. *Id.* at 1073.

232. *Id.* at 1074.

and place of business of the parties; and 3) the place where the relationship between the parties is centered.[233]

In applying *Hubbard* to *Schalliol*, the court decided that Indiana law must apply to all aspects of plaintiffs' claims against the United States for the negligent actions of their air traffic controllers sitting in Indianapolis.[234] The court reasoned that the place of the accident, Kentucky, bore little connection to the legal action concerning plaintiffs' claims against the United States' air traffic controllers.[235] Turning to the "place where the conduct causing the injury occurred," the court decided that the negligent actions of the air traffic controllers in Indiana most directly affected the aircraft's approach and subsequent destruction.[236] The court then admitted that the domiciles of the majority of the decedents were in Pennsylvania, but refused to apply Pennsylvania law to compensatory damages.[237]

However, on appeal the Third Circuit found that "*Hubbard* gives no guidance as to which factor is most important or how to 'break a tie,' so any decision by this court on which substantive law Indiana would apply would be little more than a guess."[238] In applying the *Hubbard* factors, the appellate court determined that "the first points to Indiana, the second to Pennsylvania, and the third is indeterminate."[239] The appellate court, therefore, certified questions for the Indiana Supreme Court to answer concerning the clarification of Indiana's choice-of-law rule.[240]

Schalliol presents a situation where depecage must be applied to reach a logical solution. Without its application the result creates a legal impossibility. Indiana law cannot apply to all aspects of the plaintiffs claims against the United States because to allow it would create an irreconcilable conflict, with the entry of a final judgment. All parties, including the United States, agreed that in Mary Schalliol's case, on behalf of her deceased husband, the law of Indiana applies, while against Fare and Hart Delaware, the law of Pennsylvania applies.[241] As a result of this agreement, it will be necessary to instruct a jury on the measure of damages in Pennsylvania, as well as an instruction on who, under Pennsylvania law, may recover for the wrongful death of Dennis Schalliol. Similarly, the jury will need to know how to apportion those damages to the different parties. In regards to Fare/Hart, the jury would use Pennsylvania's law of comparative fault coupled with joint and several liability.[242]

At the same time that the jury is deciding Schalliol's claims against Fare and

---

233. *See* RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 145(2) (1971).
234. *Schalliol*, 206 F. Supp. 2d at 700.
235. *Id.*
236. *Id.*
237. *Id.*
238. Simon v. United States, 341 F.3d 193, 205 (3d Cir. 2003).
239. *Id.*
240. *Id.*
241. *See supra* note 222.
242. *See* 42 PA. CONST. STAT. ANN. §§ 8322, 8324 (2002).

Hart Delaware, the judge as the trier of fact[243] will consider Indiana's laws for the measure of damages, who may recover, and the apportionment of such damages. The judge will use Indiana's rule of comparative fault (with no joint and several liability).[244] Even though the jury is only determining the liability of Fare and Hart Delaware, their decision under Pennsylvania law as to what percentage Fare and Hart are at fault will automatically determine the fault of the United States. For example, if the jury finds Fare/Hart ten percent at fault, the fault of the United States is then ninety percent. This is due to the fact that there is no question of contributory negligence on the part of Dennis Schalliol, an innocent passenger. The total fault between the defendants must equal 100%.

Without the use of depecage, an irreconcilable conflict exists between Pennsylvania and Indiana law as to whether the adult, independent children of Dennis Schalliol may recover damages. Under Pennsylvania law, adult independent children may recover,[245] while under Indiana's wrongful death act, only dependent children may recover.[246] Thus, the problem of applying Indiana law as to all issues surround the United States and Pennsylvania law as to all issues surrounding Hart Delaware and the Fare Estate is two tiered. First, Pennsylvania permits joint and several liability while Indiana does not.[247] Secondly, Pennsylvania permits different elements of damages and allows different people to recover than is allowed under Indiana law.[248]

If the district court's ruling in *Schalliol* is followed, applying Indiana law to all issues against the United States, then there is a possibility for plaintiff Schalliol to either be over or under compensated.[249] Assume that under Pennsylvania law, the jury returns a $5 million verdict broken down as follows: $4 million to Mary Schalliol; $1 million awarded to decedent Dennis Schalliol's adult independent children; and fault allocation of ninety percent to the United States and ten percent to Fare/Hart Delaware.

If the judge adopts the jury's findings concerning damages and fault apportionment, overcompensation occurs when the award to the adult independent children of Dennis Schalliol is taken into consideration. In the hypothetical, Dennis Schalliol's children are entitled to receive $1 million under

---

243. *See* 28 U.S.C. § 2402 (2002); Hamm v. Nastka Barriers, Inc., 166 F.R.D. 1, 2 (D.C. 1996).

244. IND. CODE §§ 34-51-2-8, 34-51-2-12 (2002).

245. 42 PA. CONST. STAT. ANN. § 8301 (2002); Burchfield v. M.H.M. P'ship, 43 Pa. D. & C.4th 533, 542-43 (C.P. 1999).

246. IND. CODE § 34-23-1-1 (2002).

247. *Compare* 42 PA. CONST. STAT. ANN. §§ 8322, 8324 (2002), *with* IND. CODE §§ 34-51-2-8, 34-51-2-12 (2002).

248. *Compare* 42 PA. CONST. STAT. ANN. § 8301 (2002); *Burchfield*, 43 Pa. D. & C.4th at 542-43, *with* IND. CODE § 34-23-1-1 (2002).

249. Similar problems exist between Pennsylvania's allowance of an award for conscious pain and suffering and pre-impact fear, and Indiana's bar of such awards. *Compare* IND. CODE § 34-9-3-4 *and* Estate of Sears v. Griffin, 771 N.E.2d 1136, 1138 (Ind. 2002), *with* Williams v. S.E. Pa. Transp. Auth., 741 A.2d 848, 859 (1999).

Pennsylvania law, but not under Indiana law. Therefore, the children would not be able to recover any of their award from the United States. Nevertheless, because Pennsylvania allows joint and several liability,[250] the Schalliol children would still be able to collect the full $1 million from Fare/Hart Delaware. Thus, if under Pennsylvania law, the Schalliol children collect the full $1 million from Fare/Hart, the total award jumps to $5.5 million ($1 million to Schalliol children awarded by jury plus $4.5 million to Mary Schalliol awarded by judge), while both the judge and jury found the total award to be only $5 million.

One possible solution to this conundrum is for the judge to only award Fare/Hart Delaware's portion of damages as determined by the jury. Thus, under the hypothetical posed above, the most that the Schalliol children could receive from Fare/Hart Delaware is $500,000 (ten percent of $5 million). This would solve the problem of the excessive total award as it would equal $5 million ($500,000 plus $4.5 million). However, now there exists a problem of under compensation. The jury awarded the Schalliol children $1 million, but they are only able to collect half of that amount. The Schalliol children should be able to recover their award of $1 million from either joint tortfeasor, the United States or Fare/Hart Delaware. The essence of joint and several liability is that a plaintiff is allowed to obtain all or part of a judgment from any defendant.[251] Therefore, Pennsylvania requires that plaintiffs be allowed to collect any and/or all damages from any joint tortfeasor. Applying joint and several liability to less than all defendants severely undermines Pennsylvania's laws and interests and significantly burdens either the plaintiff or those joint tortfeasors who are subject to joint and several liability.

What then is the solution to such irreconcilable conflict between two states laws? Depecage is the answer. Instead of applying Indiana law to all issues regarding Schalliol's claims against the United States, the court should have separated out three issues: 1) negligence; 2) joint and several liability; and 3) compensatory damages. Under a Restatement section 145 analysis, Indiana's elements of negligence should apply to the United States since Indiana is where the most direct cause of the aircraft's fatal crash occurred. As to issues of compensatory damages, the law of the decedent's domicile, Pennsylvania, should govern. The issue of joint and several liability is a bit trickier. Joint and several liability is a common law negligence principle; however, it has a direct correlation to the apportionment of damages.

Indiana has an interest in applying its law to protect those doing business within its borders from the daunting concept of joint and several liability, while Pennsylvania has an interest in seeing its residents fully compensated. The question is which state's law should yield? Several arguments could be made about whose law should apply to this issue. Some of those arguments include: 1) apply Indiana law to the issue of joint and several liability because the court is using Indiana's choice-of-law rule; 2) apply Pennsylvania law because the

---

250. 42 PA. CONST. STAT. ANN. §§ 8322, 8324 (2002).

251. *See* L.B. Foster Co. v. Charles Caracciolo Steel & Metal Yard, Inc., 777 A.2d 1090, 1095 (Pa. Super. 2001).

forum's law should prevail in a true conflict; 3) apply Pennsylvania law because Indiana's interest in seeing its law apply to a federal entity operating within its boarders is not as strong as the Pennsylvania's interest in seeing its law applied to its residents. To argue within the rules, any one of these arguments may be offered by the parties, but there is an argument to be made which does not fall within the normal interest analysis framework. For the law to function, Pennsylvania law must apply to the United States on the issue of joint and several liability. If it does not apply, then the court is faced with the irreconcilable conflict posed above.

This functional analysis has been used sparsely by other courts[252] and should be given consideration in the unique situation presented in *Schalliol*. Such a test should consist of factors such as ensuring "certainty, predictability and uniformity of result" and "ease of application of law to be applied . . . ."[253] A functional analysis of the law may in some instances "break the tie" between two states' equal interests in seeing their own law applied. In *Schalliol*, for the court to be able to enter a final judgment that fairly compensates parties and respects state interests, Pennsylvania law must be applied to the issue of joint and several liability.

Regardless of the approach used in deciding which state's law to apply when a true conflict exists, depecage is still an essential element. Depecage serves to minimize true conflicts by separating out and focusing in on only those issues where a true conflict exists. In examining *Schalliol*, it would be easy to say that a true conflict exists between the law of Indiana and Pennsylvania. However, a true conflict does not exist between every aspect of Indiana and Pennsylvania law. Depecage dissects state law to determine exactly where the true conflict exists.

Depecage may serve its purpose of limiting the number of true conflicts between state laws in many different settings. Although courts justify their use of depecage as a means of best respecting the interests of all states in the resolution of a multi-state controversy, permitting them to give effect to the interests of more than one state in each case, depecage does not necessarily have to coincide with interest analysis.[254] Depecage can be more than just a part of interest analysis because its true utility lies in its ability to lessen conflicts through separating out issues. So long as a court has not adopted a uniform rule that the law of the state where the tort occurred governs or that the law of the forum must govern, depecage is adaptable to any choice-of-law analysis.

Depecage is most often applied in complex situations, but arguably should be used in simple cases as well. Whether for ordinary or complex cases, the choice-of-law rules used to define substantive rights should be the same.[255] For example, depecage may be applied in marriage and divorce recognition. It is realistic to construct a coherent approach to choice of law in the area of validity

---

252. Clawans v. United States, 75 F. Supp. 2d 368, 375 (D. N.J. 1999).

253. *Id.*

254. Reavley & Wesevich, *supra* note 141, at 50 n.184.

255. Larry Kramer, *Choice of Law in Complex Litigation*, 71 N.Y.U. L. REV. 547, 549 (1996).

of marriage through depecage, splitting the traditional category into smaller units, each restricted to genuinely related issues. "It has been propounded that for each delineated category there needs to be a separate connecting factor designed to make a policy-based link between all the related issues in the category and the abstractly defined state whose law is to apply."[256]

Cases do not have to fall into the typical complex fact pattern for depecage to be applied. In *Ruiz v. Blentech Corp.*, the court applied depecage to a situation where only two states and two parties were involved.[257] The plaintiff, a citizen of Illinois, suffered an injury in his home state from an allegedly defective product manufactured in California by a California corporation. The manufacturer had dissolved, but another California corporation followed in its footsteps by purchasing its principal assets and continuing its business. The plaintiff sought recovery under a successor corporation theory, with Illinois and California differing on the rules for determining when one corporation is responsible, as a successor, for the tort liabilities of its predecessors. The court used depecage to separate the issues of successor liability and tort liability and ruled that California law would apply to the issue of corporate successor liability while Illinois law would govern the actual tort.[258]

A broader use of depecage with adequate recognition for the legal analysis being performed would help to clear up the feared complexity that depecage brings to the table. As a result, courts might realize the benefits that come from embracing complexity through the use of depecage.

CONCLUSION

An incredible shift in choice-of-law analysis has taken place over the past one-hundred years. Many great legal minds have tried to develop a conflicts of law theory that avoids the arbitrary nature of *lex loci delicti* when choosing whose law to apply. In foregoing *lex loci delicti* and its process void of any legal analysis, courts have opened themselves up to complexity. Counting contacts, considering governmental interests, and determining reasonable expectations call for complicated legal analysis by courts. While at first glance, it appears that depecage adds to the complexity, in reality depecage allows courts to limit the number of true conflicts among two differing bodies of law. Instead of laying a blanket over a multi-faceted choice-of-law problem, depecage dissects the problem eliminating issues presenting false conflicts and leaving only those issues where true conflicts exist. Thus, justice, fairness, and efficiency in achieving the best practical result are attained.

Depecage is not only a helpful tool in choice-of-law analysis, it may be an essential tool to resolving what seems to be a legal impossibility. Although

---

256. Alan Reed, *Essential Validity of Marriage: The Application of Interest Analysis and Depecage to Anglo-American Choice of Law Rules*, 20 N.Y.L. SCH. J. INT'L & COMP. L. 387, 423-24 (2000).

257. Ruiz v. Blentech Corp., 89 F.3d 320, 325-26 (7th Cir. 1996).

258. *Id.*

consisting of many different parts, our legal system is designed to function as a whole. When bodies of law conflict in such a way as to threaten this functionality, a solution must be found. Through the separation of legal issues, depecage allows entirely different bodies of law to function while still honoring the interests of each.

United States Postal Service

Statement of Ownership, Management, and Circulation

1. Publication Title	2. Publication Number	3. Filing Date
Indiana Law Review	0 0 9 0 - 4 1 9 8	9-25-03
4. Issue Frequency	5. Number of Issues Published Annually	6. Annual Subscription Price
Quarterly	Four	$30.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*

530 W. New York St., Indianapolis, Marion, Indiana 46202-3225

Contact Person: Chris Paynter

Telephone: 317-274-4440

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

530 W. New York St., Indianapolis, Marion, Indiana 46202-3225

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*

Indiana University School of Law - Indianapolis
530 W. New York St.
Indianapolis, Indiana 46202-3225

Editor *(Name and complete mailing address)*

Stephen M. Worth
530 W. New York St.
Indianapolis, Indiana 46202-3225

Managing Editor *(Name and complete mailing address)*

Jeffrey S. Beck
530 W. New York St.
Indianapolis, Indiana 46202-3225

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
Indiana University Board of Trustees	Indiana University - Bloomington
	Indiana Memorial Union Room 005
	Bloomington, Indiana 47405

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address
NONE	NONE

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*

The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:

☒ Has Not Changed During Preceding 12 Months

☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

13. Publication Title			14. Issue Date for Circulation Data Below: September 2003	
15. Extent and Nature of Circulation			Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*			891	1385 Issue 36:4
b. Paid and/or Requested Circulation	(1)	Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	718	725
	(2)	Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)*	-0-	-0-
	(3)	Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	145	550
	(4)	Other Classes Mailed Through the USPS	20	25
c. Total Paid and/or Requested Circulation *(Sum of 15b. (1), (2),(3),and (4))* ▶			883	1300
d. Free Distribution by Mail *(Samples, complimentary, and other free)*	(1)	Outside-County as Stated on Form 3541	100	50
	(2)	In-County as Stated on Form 3541	-0-	-0-
	(3)	Other Classes Mailed Through the USPS	15	10
e. Free Distribution Outside the Mail *(Carriers or other means)*			150	100
f. Total Free Distribution *(Sum of 15d. and 15e.)* ▶			265	160
g. Total Distribution *(Sum of 15c. and 15f)* ▶			1148	1460
h. Copies not Distributed			35	20
i. Total *(Sum of 15g. and h.)* ▶			1183	1480
j. Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*			75%	88%

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the 37:1 issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner: Editorial Specialist/Business Manager — Date: 9-25-03

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.